



УДК 517.1:519.6

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ДОСЛІДЖЕННЯ АДЕКВАТНОСТІ РЕГРЕСІЙНИХ МОДЕЛЕЙ МЕТОДОМ ВДОСКОНАЛЕНОЇ СЕРІЙНОЇ КОРЕЛЯЦІЇ

Студ. І.І. Коваленко, гр. МгІТ 1-18
Науковий керівник проф. С.М. Краснитський
Київський національний університет технологій та дизайну

Мета і завдання. Мета полягає в удосконаленні програмного забезпечення, що призначено для дослідження якості регресійних моделей методом серійної кореляції залишків.

Завдання полягає в розробці комп'ютерної програми, що реалізує зазначений метод з урахуванням його вдосконаленого варіанту порівняно із більш ранніми рекомендаціями (див., наприклад, [1]).

Об'єкт та предмет дослідження. Об'єктом дослідження є метод серійної кореляції дослідження якості регресивних моделей при описанні стану складних систем. Предметом дослідження є алгоритм перевірки якості регресійних моделей за допомогою зазначеного методу.

Методи та засоби дослідження. Теоретичною основою при вирішенні науково-технічної проблеми є праці провідних вчених в галузях регресійного аналізу, дослідження операцій, математичного моделювання, програмного забезпечення САПР. У теоретичних дослідженнях використано методи наближених обчислень, статистичного аналізу даних, теорії алгоритмів[1 – 2].

Наукова новизна та практичне значення отриманих результатів. Стандартне статистичне і загально математичне забезпечення не враховує більш пізні уточнення у послідовності та змісті відповідної статистичної процедури.

Результати дослідження. У регресійних дослідженнях ми, як правило, припускаємо, що помилки спостережень попарно некорельовані. Якби це припущення виявилось фактично порушеним, то можна було б очікувати, що графік залежності залишків від часу або який-небудь інший, чутливий до порядку графік, який ми б збудували з урахуванням наших обставин, допоможе встановити цей факт. Звичайно, помилки можуть корелювати найрізноманітнішими способами. Але, як правило, все зводиться до того, що вони можуть проявити серіальну кореляцію. Таку кореляцію, де залежність між помилками, віддаленими одна від одної на кроків, завжди залишається однаковою. У тих окремих випадках, коли залишки проявляють локально позитивну серіальну кореляцію, сусідні залишки у часовій послідовності стають більш схожими один на одного, ніж на інші залишки. Кореляція між залишками, які розділяє один крок (або два, або три) називається серіальною кореляцією з одиничним зсувом.

Можуть зустрітися і негативні серійні кореляції між сусідніми залишками. Одна з причин цього явища, яке зустрічається в періодичних або циклічних процесах, відома як транспортне запізнювання. Ось як це може відбуватися. Покладемо, що деяка конкретна порція продукту, введена в процес, перероблена в повному обсязі, оскільки частина продукту залишилася в трубопроводах і насосах, що забезпечують даний реактор. Тоді зафіксований вихід продукту для цієї порції виявився б незвичайно низьким. Зате наступна порція сприяла б переробці цього запасеного матеріалу, даючи тим самим незвично високий вихід.

Основний метод аналізу, яким можна скористатися при наявності серійної кореляції залишків, - це зважений метод найменших квадратів. А головна проблема в його застосуванні – пошук «інформаційної» матриці V . Матриця V являється діагональною матрицею з нерівними елементами. У деяких ситуаціях може виявитися також, що позадіагональні елементи матриці, не є рівними нулю, тобто, спостереження корельовані. Коли спостереження впорядковані в часі, елементом матриці V служитиме p_l , $l = |i -$



j причому $p_0 = 1$. Для оцінки p_t нам треба зрушувати спостереження l на кроків і обчислювати коефіцієнт кореляції за формулою, відкидаючи ті спостереження, для яких при зсуві бракуватиме пари. Отримані таким чином оцінки підставляються в матрицю V , що дає матрицю \hat{V} . Для аналізу залишків такої зваженої моделі нам потрібні оцінки матриці

$$f = P^{-1}e$$

Ось ці оцінки: $\hat{f} = \hat{P}^{-1}(Y - \hat{Y})$, де

$$\hat{P}'\hat{P} = \hat{V}$$

\hat{Y} - передбачення зваженого методу найменших квадратів,

$$\hat{Y} = X(X'\hat{V}^{-1})^{-1}X'\hat{V}^{-1}Y$$

Іншими словами, матриця \hat{f} задається наступною рівністю

$$\hat{f} = \hat{P}^{-1} \left| I - X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1} \right| Y$$

Існують два добре відомі способу перевірки того, чи є в залишках ознаки серіальної кореляції. Це критерій серій і критерій Дарбіна-Уотсона.

При відомій часовій послідовності множин залишків іноді вдається спостерігати незвичайні скупчення позитивних або негативних залишків. Візьмемо для прикладу крайній випадок. Якщо тимчасова послідовність з тридцяти залишків містить спочатку десять негативних, а потім двадцять позитивних значень, то ми маємо право очікувати, що є якась не приймалася до уваги змінна, яка змінила рівень між десятим і одинадцятим дослідом. Це означає, ми можемо досліджувати причини, що викликають подібну поведінку. Коли зустрічається така послідовність знаків, корисно мати метод, що дозволяє виносити рішення про відхилення чергування знаків в послідовності дослідів від випадкового, тобто. Виносити рішення про «ненормальність» послідовності. Припустимо, ми маємо наступну послідовність знаків:

+ + - + - - - - + + - + + +

Це знаки залишків у часовій послідовності (якими ми скористаємося). Знаки «плюс» і «мінус» можуть означати також «чоловічий» і «жіночий», «голови» і «хвости», «краще» і «гірше», «обробка А» і «обробка В» або два рівня будь-який інший дихотомічної класифікації. Припустимо, що тут в знаків n , з яких n_1 знаків плюс і n_2 знаків мінус, а u - число серій. В наведеному прикладі $n_1 = 8$, $n_2 = 6$, $u = 7$ серій, показаних круглими дужками:

(++)(-)(+)(- - - -)(++)(-)(+ + +)

Є широко поширений критерій для виявлення серіальної кореляції певного виду, який називається критерієм Дарбіна-Уотсона. За допомогою критерію Дарбіна-Уотсона можна перевірити нуль-гіпотезу H_0 про те, що все проти альтернативи

$$H_1: p_s = p^s$$

Висновки. Отримано алгоритм та програмне забезпечення для дослідження регресійних моделей методом вдосконаленої серійної кореляції залишків. **Ключові слова:** програмне забезпечення, регресійні моделі, метод серійної кореляції залишків.

ЛІТЕРАТУРА

1. Дрейпер Н. Р. Прикладной регрессионный анализ: монографія / Дрейпер Н. Р., Смит Г. — Москва, Санкт-Петербург, Киев: Диалектика, Вильямс, 2016. -910 с.
2. Краснитський С.М. Векторні випадкові величини і випадкові процеси / С.М. Краснитський., В.Ю.Щербань та ін. – К.:Конус-Ю, 2008.- 191 с.
3. Щербань В.Ю. Математичні моделі в САПР. Обрані розділи та приклади застосування / В.Ю.Щербань, С.М.Краснитський, В.Г.Резанова – К.:КНУТД, 2010.-220 с.