

ИСПОЛЬЗОВАНИЕ МОДЕЛИ ARIMA ДЛЯ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ ЗАБОЛЕВАЕМОСТИ САЛЬМОНЕЛЛЕЗОМ В ХАРЬКОВСКОЙ ОБЛАСТИ

Богданов Сергей Александрович, студент группы 345

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

В работе использован метод авторегрессионного интегрированного скользящего среднего (autoregressive integrated moving average, ARIMA), который является расширением модели скользящего среднего (ARMA). Эти модели используются при работе с временными рядами для более глубокого понимания данных или предсказания будущих точек ряда, либо для прогнозирования будущих точек в ряду (прогнозирование), с такими коэффициентами:

p: как порядок отставания (lag order).

d: степень различия (degree of differencing)

q: порядок скользящего среднего (moving average)

Модель ARIMA (p, d, q). – одна из наиболее популярных моделей для построения краткосрочных прогнозов, получается путем интегрирования ARMA(p,q) и задаётся следующим образом:

$$\left(1 - \sum_{i=1}^p \Phi_i * L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i * L^i\right) \epsilon_t \quad (1)$$

где X_t – временной ряд вещественных чисел;

t – целый индекс;

L – оператор задержки;

Φ_i – параметры авторегрессионной части модели;

θ_i – параметры скользящего среднего;

ϵ_t – значения ошибки;

d – положительное целое, задающее уровень дифференцирования.

В рамках исследования проведен анализ заболеваемости сальмонеллезом в Харьковской области, данные были проверены на наличие тренда, сезонности, применен тест Дики–Фуллера, для всех выборок исследованы предсказание заболевания по возрасту, половым признакам и для популяции в целом.

Проведен анализ заболеваемости сальмонеллезом в Харьковской области в 2015–2017. Прогнозирование временного ряда показало высокую точность на короткий промежуток времени по критериям возраста и пола.

**Научный руководитель – Чумаченко Д.И., к.т.н., доцент кафедры математического моделирования и искусственного интеллекта*