

ПРИМЕНЕНИЕ МЕТОДОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ ДЛЯ ЭКОЛОГИЧЕСКОГО КОНТРОЛЯ

Н.С. Завизена, С.А. Семериков

Построение статистических экологических моделей позволяет нам произвести регрессионный анализ, оценивающий динамику обрабатываемых данных. При этом многофакторные наблюдения приходится накапливать на достаточно длительном промежутке времени, так как для устойчивости регрессионной модели в простейшем линейном случае количество наблюдений должно хотя бы в два раза превышать количество факторов наблюдения; усложнение модели (параметризация факторов) приводит к росту минимального необходимого количества наблюдений, что далеко не всегда возможно.

Другим фактором, сдерживающим в отдельных случаях применение регрессии, является то, что определяемая в ходе анализа функция регрессии лишь формально устанавливает соответствие между результирующими и объясняющими переменными, хотя они в действительности могут и не состоять в причинно-следственных отношениях. Поэтому устанавливаемые в ходе регрессионного анализа связи могут иногда истолковываться как причинно-следственные. Таким образом, могут возникнуть так называемые нонсенс-регрессии, не имеющие практического значения. Поэтому перед применением аппарата регрессии необходимо решить, какой из факторов рассматривать как результирующий, а какие из регистрируемых величин - как объясняющие.

Для решения этой задачи необходимо решить некоторые частные задачи, среди которых можно выделить следующие:

1. Анализ, отбраковка и восстановление аномальных (сбитых) или пропущенных измерений. Эта задача связана с тем, что исходная экспериментальная информация обычно неоднородна по качеству. Наряду с основной массой результатов прямых измерений, получаемых с возможно малыми погрешностями, в экспериментальных данных часто имеются грубые промахи, вызванные просчётами наблюдателя, сбоями при вводе в ЭВМ, аномалиями в работе измерительных приборов и т.д. Без анализа качества данных, устранения или уменьшения влияния аномальных данных на результаты последующей обработки можно прийти к ложным выводам об изучаемом объекте или явлении.

2. Экспериментальная проверка законов распределения экспериментальных данных, оценка параметров и числовых характеристик наблюдаемых величин или процессов. Выбор методов последующей обработки, направленной на построение и проверку адекватности математической модели исследуемому явлению, существенно зависит от закона распределения наблюдаемых величин. Так, при использовании для обработки процедур классического регрессионного анализа, в первую очередь, необходимо выяснить, является ли закон распределения наблюдаемых величин гауссовским и некоррелированным и т.д. Получаемые при решении этой задачи выводы о природе экспериментальных данных могут быть как весьма общими (независимость измерений, их равноточность, аддитивный характер погрешностей и др.), так и содержать детальную

информацию о статистических свойствах данных (вид закона распределения, его параметры). Решение центральной задачи предварительной обработки не является чисто математическим, а требует также содержательного анализа изучаемого процесса, схемы и методики проведения эксперимента.

3. ***Сжатие и группировка исходной информации при большом объёме экспериментальных данных.*** При этом должны быть учтены особенности их законов распределения, выявленные на предыдущем этапе обработки.
4. ***Объединение нескольких групп измерений, полученных, возможно, в различное время или в различных условиях, для совместной обработки.***
5. ***Выявление статистических связей и взаимовлияния различных измеряемых факторов и результирующих переменных, последовательных измерений одних и тех же величин.*** Решение этой задачи позволяет отобрать те переменные, которые оказывают наиболее сильное влияние на результирующий признак. Выделенные факторы используются для дальнейшей обработки методами регрессионного анализа. Анализ корреляционных связей делает возможным выдвижение гипотез о структуре взаимосвязи переменных и, в конечном итоге, о структуре модели явления.

В ходе предварительной обработки кроме указанных часто решаются и другие задачи, имеющие частный характер: отображение, преобразование и унификацию типа наблюдений, визуализацию многомерных данных и др.

Следует отметить, что в зависимости от конечных целей исследования, сложности изучаемого явления и уровня априорной информации о нём объём задач, выпол-

няемых в ходе предварительной обработки, может существенно изменяться. То же самое можно сказать и о соотношении целей и задач, которые решаются при предварительной обработке и на последующих этапах статистического анализа, направленных на построение модели явления. Так, например, если целью эксперимента является измерение значения неизвестной, но заведомо постоянной величины путём прямых многократных измерений с помощью средства измерений с известными характеристиками погрешностей, то полная обработка результатов измерения ограничивается простейшей предварительной обработкой данных (оценкой математического ожидания). В то же время, если измеряемая величина является переменной, а закон распределения погрешностей измерительного прибора неизвестен, то для решения конечной задачи потребуются проведение как предварительной обработки данных, так и применение статистических методов исследования физических зависимостей.

Для решения задач предварительной обработки используются различные статистические методы: проверка гипотез, оценивание параметров и числовых характеристик случайных величин и процессов, корреляционный и дисперсионный анализ. Для предварительной обработки, оказывающей, как следует из сказанного, первостепенное влияние на качество решения конечных задач исследования, характерно информационное решение основных задач, когда повторно возвращаются к решению той или иной задачи после получения результатов на последующем этапе обработки.