

АБУСЕВ Р. А., ЛУМЕЛЬСКИЙ Я. П.

**СТАТИСТИЧЕСКИЕ МОДЕЛИ КЛАССИФИКАЦИИ
МНОГОМЕРНЫХ НАБЛЮДЕНИЙ**

Введение

Задачи поточечной и групповой классификации для параметрических семейств распределений занимают важное место в прикладном статистическом анализе (см., например, [3], [16], [17], [18]). Причисление одного объекта, описываемого случайным вектором (случайной матрицей), к одному из двух или более классов, заданных обучающими выборками, носит название поточечной классификации (см. [16], [29], [32]). При групповой статистической классификации к одному из нескольких классов, определенных как и выше, относится n_0 таких объектов ($n_0 > 1$). Решение задач статистической классификации имеет важное теоретическое значение и многочисленные приложения: в технической и медицинской диагностике, типологии социально-экономических явлений, при анализе экологической информации.

Опишем одну задачу, где успешно используются методы поточечной и групповой статистической классификации. Плохое состояние экологической обстановки вызывает многие заболевания, которые носят название экopatологии. Пусть имеются M различных экopatологий на данной территории, заданных обучающими выборками π_{i0} , $i = 1 \div M$, с k медицинскими диагностическими показателями. На конкретной территории имеется группа n_0 больных, работающих на одном предприятии. Требуется определить тип экopatологии.

При $n_0 = 1$ речь идет об отнесении конкретного больного к одному из типов экopatологий, а при $n_0 > 1$ возникает задача групповой классификации. При $n_0 > 1$ появляется возможность более обоснованно судить о влиянии ухудшившейся экологической обстановки на здоровье работников предприятия. При решении ряда задач методы групповой классификации позволяют в 5–7 раз уменьшить вероятность ошибок, поскольку рассматриваются группы больных в 30–45 человек.

Настоящая работа, которая носит, в основном, обзорный характер, посвящена построению асимптотически оптимальных решающих правил поточечной и групповой статистической классификации в слу-

чае многомерного нормального распределения и распределения Уишарта. Особое внимание при построении решающих правил классификации уделяется использованию несмещенных и байесовских оценок для плотностей распределений. Состоятельность указанных выше оценок обеспечивает асимптотическую оптимальность классификационных тестов. Впервые найдены также несмещенные и байесовские оценки для плотностей распределений достаточных статистик в задачах групповой классификации многомерных наблюдений. Для описания свойств различных состоятельных оценок плотностей распределений для многомерных параметрических семейств находятся квадратические погрешности этих оценок.

Настоящая статья отражает основные результаты исследований, проведенных в Пермском государственном университете за последние 20 лет. Эти результаты публиковались в работах [1], [3], [7], [14], [23] и др. Ниже приводятся также новые результаты, относящиеся к оцениванию и классификации в случае распределения Уишарта.