

**С. В. Хасанов** (Уфа, УГАТУ). Методы межсессионного контроля знаний по математике.

Задача исследования зависимости успеваемости студентов от нескольких факторов может решаться методами множественного регрессионного анализа. В работе, представленной данным сообщением, рассматриваются способы отбора наиболее существенных из этих факторов.

Пусть, например, имеются данные о зависимости случайной величины  $Y$  — результатов сессии от результатов межсессионного контроля:  $X_1$  — результатов коллоквиума и  $X_2$  — результатов контрольной работы (использовались данные первого семестра в процентах). Стандартными методами получено уравнение множественной регрессии:  $Y = 1,667 + 0,654X_1 + 0,239X_2$ .

Коэффициенты при объясняющих переменных показывают, что результаты коллоквиума сильнее влияют на итоговую успеваемость. Если не учитывать переменную  $X_2$ , парное уравнение регрессии  $Y$  по  $X_1$ , имеет вид:  $Y = 1,466 + 0,837X_1$ . Коэффициент при  $X_1$  здесь выше, так как он учитывает воздействие на  $Y$  не только переменной  $X_1$ , но и косвенно корреляционно связанной с ней переменной  $X_2$ .

При использовании этой модели на практике возникают некоторые проблемы. К их числу относится, например, мультиколлинеарность т. е. высокая взаимная коррелированность объясняющих переменных. Уравнения регрессии в этом случае, как правило, не имеют реального смысла.

Одним из методов уменьшения мультиколлинеарности является использование пошаговых процедур отбора наиболее информативных переменных. На первом шаге рассматривается лишь одна объясняющая переменная, имеющая с  $Y$  наибольший коэффициент детерминации. На втором шаге включается в регрессию новая объясняющая переменная, которая в паре с предыдущей имеет с  $Y$  наиболее высокий коэффициент детерминации и т. д. до тех пор, пока будет увеличиваться (скорректированный) коэффициент детерминации.

В работе исследовалась зависимость результатов  $Y$  сессии от результатов  $X_1$  контрольной работы,  $X_2$  коллоквиума и  $X_3$  проверок расчетно-графических работ (в процентах по 7 группам). Было получено уравнение регрессии:  $Y = -11,812 + 0,1X_1 + 0,553X_2 + 0,535X_3$ . После применения пошагового отбора информативных переменных это уравнение приняло вид:  $Y = -19,221 + 0,572X_2 + 0,537X_3$ . Значимость уравнения по  $F$ -критерию (на уровне 0,95) повысилась с  $F = 34,2$  до  $F = 50,78$ .

#### СПИСОК ЛИТЕРАТУРЫ

1. Хасанов С. В. Корреляционно-регрессионный анализ в задачах нефтегазового дела. Уфа: УГНТУ, 2005, 82 с.