

В.Т. Олейников, А.А. Мосягин
МЕТОДЫ КОМПЬЮТЕРНОГО АНАЛИЗА ОБСТАНОВКИ
НА ПОТЕНЦИАЛЬНО ОПАСНЫХ ОБЪЕКТАХ

Прогнозирование пожаров и ЧС на потенциально опасных объектах предлагается осуществлять на основе методов интеллектуального анализа данных (ИАД), который заключается в применении алгоритмов обработки данных для выявления скрытых тенденций, закономерностей, взаимосвязей и перспектив развития процесса, учет которых помогает повысить качество принимаемых решений. Все методы, используемые в настоящее время в интеллектуальном анализе данных, являются логическим обобщением различных аналитических подходов, большинство которых известны уже на протяжении десятилетий. Новизна ИАД заключается в расширении сферы применения этих методов в управлении.

1. **Проверка предположительных утверждений** - гипотез относительно механизмов формирования имеющихся данных.

По своему прикладному содержанию проверяются гипотезы о следующих предположениях:

- об общем виде закона распределения данных;
- об однородности нескольких групп данных;
- о числовых значениях параметров генеральной совокупности;
- о виде зависимости между факторами;
- о независимости и стационарности ряда наблюдений.

Процедура обоснованного сопоставления гипотезы H_0 с имеющимися данными x_1, \dots, x_n , осуществляется с использованием того или иного статистического критерия и называется статистической проверкой гипотез. Результат может быть отрицательным - данные противоречат гипотезе, и она отвергается, либо неотрицательным - данные не противоречат высказанной гипотезе. В последнем случае гипотеза признается правдоподобной, возможно, наряду с некоторыми другими гипотезами [2].

Опишем коротко логическую схему процедуры проверки гипотезы H_0 .

А. Выбирается некоторая функция γ^n от результатов наблюдения x_1, \dots, x_n , $\gamma^n = \gamma(x_1, x_2, \dots, x_n)$, называемая критической статистикой, которая является случайной величиной и, в предположении справедливости гипотезы H_0 , подчинена хорошо изученному затабулированному закону распределения с плотностью $f_{\gamma^n}(u)$. Содержательно эта статистика обычно определяет меру расхождения имеющихся данных с гипотезой H_0 .

В. Задается величина уровня значимости критерия α , определяющего вероятность ошибочного отклонения проверяемой гипотезы H_0 . Величина α определяется на основе субъективной оценки эксперта.

С. Из таблиц распределения $f_{\gamma^n}(u)$ находятся точка $\gamma_{\alpha/2}^{\min}$ и точка $\gamma_{\alpha/2}^{\max}$, разделяющие область значений γ^n на три части (рис. 1): область неправдоподобно малых (I), неправдоподобно больших (III) и правдоподобных значений (II).

Д. По имеющимся данным вычисляется численное значение $\gamma^n = \gamma(x_1, x_2, \dots, x_n)$. Если окажется, что полученное значение принадлежит области II, то гипотеза H_0 не противоречит исходным данным. В противном случае H_0 отвергается.

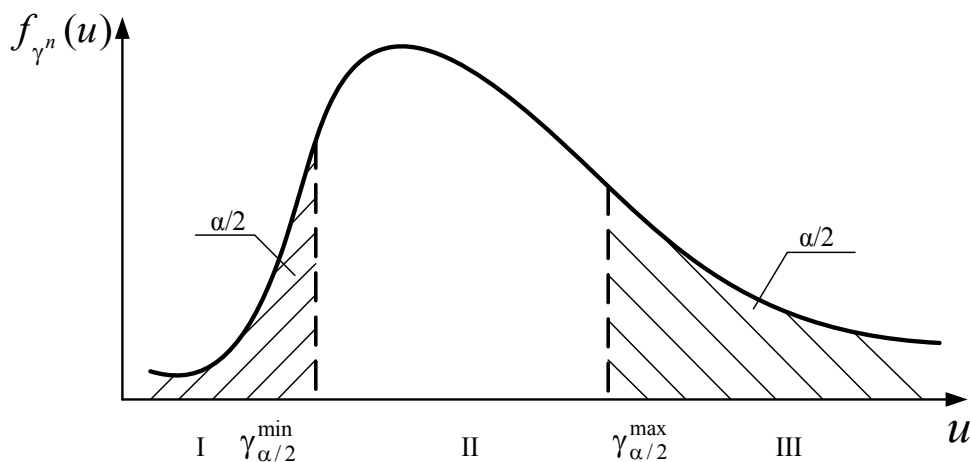


Рис. 1. Область значений соответствия гипотезе

В математической статистике исследованы многочисленные критические статистики для проверки всех типов гипотез. Широко используются t - статистика Стьюарта, F - статистика Фишера, χ^2 (хи-квадрат), ω^2 , Колмогорова и др.

Если данные извлечены из нормальных совокупностей, могут быть использованы t - и F - статистики.

2. Регрессионный анализ направлен на аппроксимацию функции и выявление зависимостей между входными и выходными переменными.

В общем виде задача аппроксимации ставится так: найти вектор-функцию:

$$f(x) = \begin{pmatrix} f_1(x) \\ \cdot \\ \cdot \\ f_k(x) \end{pmatrix}, \quad (1)$$

где $f_i(x)$ дает наилучшую аппроксимацию выходного p -мерного вектора y на исходном множестве точек $\{x_i\} i = 1, \dots, n$.

Для этого вводятся невязки $\delta_{ki} = \delta(f_k(x_i))$, характеризующие погрешность оценки результирующего признака y_k , $k = l, \dots, p$ в точке x_i , и функционал $\Delta(f) = \Delta\{\delta_{ki}\}$, $k = l, \dots, p$, $i = l, \dots, n$, определяющий меру адекватности (ошибки). Они во многих случаях определяются на основе опыта и знаний специалистов. Далее решается оптимизационная задача поиска экстремума функции $\Delta(f)$ (минимизация ошибки). В регрессионных статистических моделях функция $f(x)$ определяется как условное математическое ожидание, т.е. модель имеет вид $y = f(x) = M(y/x)$, $f(x)$ называют регрессией.

В частном случае единственного результирующего признака y - и F - класса линейных функций получаем линейное уравнение регрессии:

$$y = f(x) = M(y|x) = a_0 + \sum_{i=1}^m a_i x_i. \quad (2)$$

Дискриминантный анализ. Предполагается, что имеется совокупность объектов, разбитая на несколько групп (т.е. для каждого объекта мы можем сказать, к какой группе он относится). Для каждого из этих объектов имеются измерения нескольких количественных характеристик. Методы дискриминантного анализа позволяют отнести произвольный объект к тому или иному классу на основании измерений его количественных характеристик.

Необходимо заметить, что по методам статистического анализа создано очень большое число широко используемых пакетов программ.

Когда руководитель или эксперт оценивает результаты мониторинга, то его интересуют не только значения физических параметров, но и их тенденция, динамика развития. Система компьютерного мониторинга, анализируя полученные уравнения, может классифицировать их в соответствии с заданными типами. Тип графика (полученной аппроксимирующей кривой или прямой линии) характеризует характер процесса: его устойчивость, тенденцию, возможность резких изменений и т.п.

Таким образом, предотвращение пожаров и ЧС на потенциально опасных объектах возможно осуществлять на основе методов интеллектуального анализа данных, которые позволяют прогнозировать изменения контролируемых параметров опасного технологического процесса.

Литература

1. Трахтенгерц Э.А. Компьютерная поддержка формирования целей и стратегий. – М.: СИНТЕГ, 2005. – 224 с.
2. Брушлинский Н.Н. Математические методы и модели управления аварийно-спасательными службами // Сб. науч. трудов "Методы кибернетики и информационные технологии" РАЕН. – Саратов: Гос. унив., 1994. – С. 21-28.