

Каждому мониторинговому или укрупненному мониторинговому участку соответствует своя структура передачи информационных сигналов, а также своя проработка информации на каждом информационном уровне (от 1 до 4). Основными условиями при организации рабочего процесса замкнутой и непрерывной передачи информационных сигналов становятся согласованность работы технических средств и систем; оперативные возможности средств обработки сигналов; объемы накапливаемой и передаваемой информации должны отвечать потребностям соответствующего мониторингового участка; информационно-поисковые системы должны быть толерантными между собой.

Все это позволит обеспечить совместимость информационных потоков, необходимую плотность и периодичность оперативной информации для моделируемого и прогнозируемого управления гидроэнергетических объектов.

Б.Ю. Лемешко, С.Н. Постовалов

СИСТЕМА СТАТИСТИЧЕСКОЙ ОБРАБОТКИ СЛУЧАЙНЫХ НАБЛЮДЕНИЙ

Новосибирский государственный технический университет

Программная система, разработанная в Новосибирском государственном техническом университете [1–3], предназначена для всестороннего статистического анализа наблюдений одномерных непрерывных случайных величин. Она ориентирована на наиболее общее представление исходных наблюдений в виде *частично группированной выборки*, частными случаями которой являются негруппированные, группированные и цензурированные выборки. Система позволяет получать оценки максимального правдоподобия параметров **26** непрерывных законов распределения, проверять гипотезы о согласии по критериям χ^2 -Пирсона, отношения правдоподобия, Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса, *идентифицировать* закон распределения, выполнять группирование и сортировку вы-

борки, моделировать псевдослучайные выборки, подчиненные заданным законам распределения.

При проверке гипотез о согласии для найденного значения соответствующей статистики S^* вычисляется вероятность

$$p = P\{S > S^*\} = \int_{S^*}^{\infty} g(s) ds,$$

где $g(s)$ — плотность распределения статистики при условии истинности нулевой гипотезы. При заданном уровне значимости α гипотеза о согласии не отвергается, если $p > \alpha$. В системе используется несколько критериев и выбор наилучшего распределения осуществляется по совокупности выводов.

Для случаев, когда исходные наблюдения заданы *неточно*, разработана методика использования непараметрических критериев согласия Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса [4, 5].

Для сокращения объема хранимых и передаваемых данных и получения *робастных* оценок в системе предусмотрена процедура группирования. Наряду с обычными равномерным и равночастотным (равновероятным) группированием, *используется асимптотически оптимальное группирование данных* [2], минимизирующее потери информации о параметрах распределения при группировке. Применение асимптотически оптимального группирования обеспечивает *максимальную мощность* используемых критериев χ^2 -Пирсона и отношения правдоподобия при близких альтернативах.

Использование группирования наблюдений при оценивании параметров позволяет получать робастные оценки. Оценки, определяемые по предварительно сгруппированным данным, оказываются устойчивыми в ситуациях, когда наши предположения о наблюдаемом законе сильно отличаются от действительности. *Группирование выборки позволяет резко снизить влияние аномальных наблюдений*, а иногда и совсем исключить их влияние. Это дает возможность проводить эффективную отбраковку грубых ошибок измерений.

Отличительной особенностью *разрабатываемой* версии программной системы является возможность использования в качестве вероятностных моделей *усеченных законов распределений и смесей* распределений [6], что существенно расширяет

область применения и позволяет более точно идентифицировать закон распределения случайной величины.

В общем случае параметр смеси w двух распределений с функцией плотности $f(x) = wf_1(x) + (1-w)f_2(x)$, принадлежит интервалу $[a, b]$, где

$$a = \max_{x \in A} \frac{f_2(x)}{f_2(x) - f_1(x)} \leq 0, \quad b = \min_{x \in B} \frac{f_2(x)}{f_2(x) - f_1(x)} \geq 1,$$

где $A = \{x | f_2(x) < f_1(x)\}$, $B = \{x | f_2(x) > f_1(x)\}$,
 $a = 0$, если $\exists x: f_2(x) = 0 \wedge f_1(x) \neq 0$, $b = 1$, если $\exists x: f_1(x) = 0 \wedge f_2(x) \neq 0$. В системе предусмотрена возможность использования смесей распределений как с параметром $w \in [0, 1]$, так и $w \notin [0, 1]$.

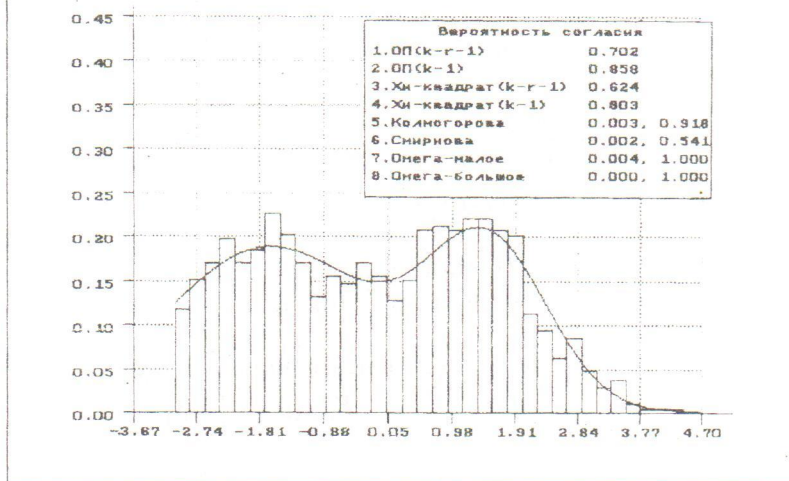
Усеченные распределения имеют плотность вида

$$f(x) = \begin{cases} f_0(x) / (F_0(\alpha) - F_0(\beta)), & x \in [\alpha, \beta], \\ 0, & x \notin [\alpha, \beta], \end{cases}$$

где α и β — параметры усечения, $f_0(x)$ — исходная функция плотности. Параметры усечения задаются пользователем или оцениваются по порядковым статистикам. На рис. приведены результаты оценивания параметров и проверки гипотез о согласии смеси двух усеченных нормальных распределений по группированной выборке объемом 1000 наблюдений. В правом верхнем углу показаны результаты проверки гипотез о согласии по перечисленным выше критериям. Так как исходная выборка группирована, то вероятность согласия по непараметрическим критериям получается в виде интервала.

Система может использоваться при проведении научных исследований в любой прикладной области, связанной с регистрацией наблюдений, а значит и с проявлением случайных закономерностей: при обработке результатов научных экспериментов, при испытаниях на надежность, при контроле качества, при метрологической аттестации измерительных приборов и обработке измерительной информации, при анализе результатов социологических обследований, при обработке наблюдений медицинских и биологических исследований, при математическом и имитационном моделировании сложных систем, анализе результатов моделирования и т. д.

Смесь: Усеченное слева (-3.0749) Нормальное с масштабом 1.3247 со сдвигом -1.5772 и Усеченное слева (-3.0749) Нормальное с масштабом 0.3006 со сдвигом 1.3000 с параметром смеси 0.5023



Оценивание параметров смеси двух усеченных нормальных распределений по группированной выборке

Созданное и развиваемое программное обеспечение может широко использоваться в учебном процессе при освоении дисциплин, связанных с обработкой информации и информационно-измерительными системами, теорией измерений, при освоении курсов теории вероятностей и математической статистики, теории случайных процессов, исследования операций, методов моделирования, статистической физики и т. д.

Система реализована на языке C++, с использованием возможностей объектно-ориентированного программирования. Библиотека классов распределений может использоваться в составе других систем. Оболочка системы базируется на библиотеке классов *Turbo Vision*.

Литература

1. Лемешко Б.Ю. Статистический анализ одномерных наблюдений случайных величин: Программная система / Новосиб. гос. техн. ун-т. — Новосибирск, 1995. 125 с.
2. Денисов В.И., Лемешко Б.Ю., Цой Е.Б. Оптимальное группирование, оценка параметров и планирование регрессионных экспериментов: В 2 ч. / Новосиб. гос. техн. ун-т. — Новосибирск, 1993. 346 с.

3. Лемешко Б.Ю., Постовалов С.Н. Статистический анализ одномерных наблюдений по частично группированным данным // Известия высших учебных заведений. Физика. — Томск, 1995. — № 9. — С. 39–45.
4. Лемешко Б.Ю., Постовалов С.Н. К использованию непараметрических критериев по частично группированным данным // Сб. научных трудов НГТУ. — 1995. — № 2. С. 21–30.
5. Лемешко Б.Ю., Постовалов С.Н. Статистический анализ наблюдений, имеющих интервальное представление // Сб. научных трудов НГТУ. — 1996. — № 1. (в печати)
6. Лемешко Б.Ю., Постовалов С.Н. Статистический анализ смесей распределений по частично группированным данным // Сб. научных трудов НГТУ. — 1995. — № 1. С. 25–31.

Ю.Т. Морозов

МЕТОДИКИ ПОСТРОЕНИЯ ПРИБЛИЖЕННЫХ ОЦЕНОК ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ СЛОЖНЫХ СИСТЕМ

Санкт-Петербургский государственный технический университет

При оценке показателей надежности и живучести структурно-сложных технических систем часто оказывается, что бинарные модели дают практически единственную возможность найти решение задачи. Это связано с высокой размерностью задачи, которая не позволяет использовать другие более сложные модели. Однако, и в случае использования бинарных моделей, для получения решения требуются значительные объемы вычислительной работы. Попытка использовать в процессе решения средства вычислительной техники также не приводит к значительному успеху. При машинных расчетах ограничения разрядной сетки недостаточно хорошие, с точки зрения точности, алгоритмы округления чисел приводят к тому, что не время решения задачи, а число арифметических операций, выполнение которых необходимо для получения решения, является важнейшим ограничивающим фактором, не позволяющим находить оценки показателей надежности с требуемой точностью.

Одним из возможных направлений снижения объема вычислительной работы при решении задач оценки показателей надежности сложных систем является переход на построение