естественными или составлять подмножества естественных языков (диалекты).

3) Могут быть введены уровни общения, характеризуемые сложностью и детализацией задания, что обусловлено различной квалификацией абонентов. В частности, если обращение в систему соответствует определенной сложности, то ответ не должен выходить из этой категории сложности. Кроме того, количество дополнительной информации, выдаваемой абоненту при ответе, будет зависеть от категории сложности.

Универсальные пред- и постпроцессоры на практике часто объединяются в общий процессор ввода — вывода. Разумеется, все компоненты системы связаны с операционной системой (ОС) банка данных, что не указано на рисунке специально. Более того, все функциональные части системы могут быть реализованы как блоки программы и являться частью ОС, или могут быть реализованы в различных блоках оборудования.

Предложенная фукциональная организация была положена в основу разработанного в Вычислительном центре СО АН СССР на базе ЭВМ БЭСМ-6 банка данных молекулярной спектроскопии. Доступ в банк обеспечивает как с помощью перфокарт в режиме пакетной обработки, так и с терминалов 3 типов: телетайпа Т-63, электрической пишущей машинки «Консул-254» и алфавитно-цифрового дисплея «Videoton — Computer-344». Терминалы связаны с ЭВМ как непосредственно, так и через телефонные линии. Дистанционная работа с удаленных терминалов обеспечивается аппаратурой передачи данных «Диалог», позволяющей войти в стандартный канал связи. Комплекс «Диалог» разработан в ВЦ СО АН СССР под руководством инженера Б. В. Морозова. Комплекс включает в себя модуляторы и демодуляторы сигналов; устройства повышения достоверности приема, использующие помехоустойчивый код с обнаружением ошибки в одном разряде плюс проверку формы импульса в каждом разряде; адаптеры для сопряжения аппаратуры передачи данных с терминалами и мультиплексорным каналом БЭСМ-6. Скорость работы 200 бодов.

В качестве ВЗУ использованы магнитные диски. Работа с терминалов реализована на основе операционной системы, разработанной в Институте прикладной математики АН СССР.

Банк химических данных на базе БЭСМ-6 находится в режиме рабочей эксплуатации. Дистанционный доступ реализован к части данных, а именно, к ИПС электронной спектроскопии с объемом около 5 тыс. документов *. Работа в этом режиме демонстрировалась в июне 1974 года на Всесоюзной конференции по автоматизации научных исследований в г. Новосибирске и на IV Международной конференции CODATA, проходившей в пос. Цахкадзор (Армянской ССР). В последнем случае связь с банком данных в г. Новосибирске осуществлялась через районную телефонную линию на участке Цахкадзор — Ереван, а затем через междугородний канал связи Ереван — Москва — Новосибирск протяженностью свыше 5000 км. Время ответа абоненту около 1 мин. Диалог с банком возможен на русском и английском языках, причем ответ дается на языке последнего обращения в систему.

Поступило в редакцию 14 ноября 1974 г.

УДК 519.21

В. М. ЕФИМОВ, З. А. ЛИВШИЦ, А. А. НЕСТЕРОВ
(Новосибирск)

ОЦЕНКА ВЕРОЯТНОСТИ ПО СОВОКУПНОСТИ МОМЕНТОВ

В ряде практических приложений определение вероятностей связано с существенными трудностями, в то время как аналитическое или экспериментальное определение моментов вполне доступно. Например, при экспериментальном исследовании контуров спектральных линий аппаратурно просто определяются моменты путем использования соответствующих масок**. Другим примером может служить вычисление вероятности для суммы случайных величин, распределение которой слабо сходится к нормальному.

Ниже дается оценка сверху для вероятности $P(x \geqslant a)$ для неотрицательных случайных величин. Полученные результаты легко переносятся на случай оценки вероятности $P(|x| \geqslant a)$.

---

* В. С. Бочкарев, Ю. П. Дробышев, В. А. Коптюг, И. К. Коробейничева, В. И. Лобанов, Р. С. Нигматуллин. Машинная информационно-поисковая система для электронной спектроскопии.— Автометрия, 1972, № 4.

** А. М. Искольдский, М. И. Кудряшов. О восстановлении оптических сигналов в исследованиях быстропротекающих процессов.— Автометрия, 1972, № 5.

I. Пусть случайная величина x принимает только положительные значения. Задана совокупность ее начальных моментов

$$\int_0^\infty f(x)\,x^k dx = C_k \qquad (k = 0,\ 1,\ \ldots,\ n). \tag{1}$$

Требуется найти оценку сверху для вероятности $P(x \geqslant a) = \int_a^\infty f(x)\,dx$. В такой постановке задача сводится к нахождению плотности вероятности $f(x)$, максимизирующей функционал

$$P(x \geqslant a) = \int_a^\infty f(x)\,dx$$

при условиях (1).

В соответствии с принципом максимума задача поиска $f(x)$ сводится к максимизации по $t(0 \leqslant t \leqslant \infty)$ функции

$$H(\alpha_k,\ x,\ f) = f\left[1\,[x-a] + \sum_{k=0}^{n} \alpha_k x^k\right],$$

где $1\ [Z] = 0$ при $Z < 0$; $1\ [Z] = 1$ при $Z \geqslant 0$. Отсюда получаем, что

$$f = \begin{cases} 0, & \text{если} \quad 1\,[x-a] + \sum_{k=0}^{n} \alpha_k x^k < 0, \\ \infty, & \text{если} \quad 1\,[x-a] + \sum_{k=0}^{n} \alpha_k x^k > 0. \end{cases}$$

Из условия нормировки $C_0 = 1$ следует, что функция f может быть неограниченной лишь в точках касания функции

$$1\,[x-a] + \sum_{k=0}^{n} \alpha_k x^k$$

оси абсцисс, т. е. решением является дискретное распределение. Структура этого распределения определяется поведением функции

$$1\,[x-a] + \sum_{k=0}^{n} \alpha_k x^k.$$

Для четного n число дельта-функций равно $n+2/2$; при этом дельта-функция с наибольшей абсциссой находится в точке $x = a$. Для нечетного n число дельта-функций равно $n+3/2$; при этом дельта-функция с минимальной абсциссой находится в точке $x = 0$ и дельта-функция с максимальной абсциссой находится в точке $x = a$.

Анализ показывает, что вероятность $P(x \geqslant a)$ является кусочно-гладкой функцией a. На участке $a_l \leqslant a \leqslant a_{l+1}$ $P(x \geqslant a) = P_l(a)$ $(l = 0, \ldots, n)$; вероятность $P_l(a)$ служит решением системы уравнений

$$\sum_{\nu=1}^{\nu=l} P_\nu x_\nu^k = C_k \qquad (k = 0,\ \ldots,\ l), \tag{2}$$

где для любых l $x_l = a$ и нечетных l $x_l = 0$. Решение системы уравнений (2) имеет следующий вид:

$$P_l(a) = \frac{\Delta_l}{\sum_{k=0}^{s} a^{l-k} \sum_{i+j=k+2} \Delta_{lij}(-1)^{i+j}}. \tag{3}$$

В (3) $s = l$ для четных l и $s = l-1$ для нечетных l; Δ_l — определитель порядка $l+2/2$ для четного l и $l+1/2$ для нечетного l, имеющий следующую структуру:

$$\Delta_l = \begin{vmatrix} C_l C_{l-1} C_{l-2} \ldots \\ C_{l-1} C_{l-2} C_{l-3} \ldots \\ C_{l-2} C_{l-3} C_{l-4} \ldots \\ \cdot \quad \cdot \quad \cdot \quad \ldots \\ \cdot \quad \cdot \quad \cdot \quad \ldots \end{vmatrix}, \tag{4}$$

Δl_{ij} — миноры определителя Δ_l.

Левая граница $a_l (l=0,\ldots, n)$ применимости оценки (3) определяется как наибольший корень уравнения

$$\Delta_l(a) = \begin{vmatrix} a^{m-1}C_l & C_{l-1} \cdots \\ a^{m-2}C_{l-1}C_{l-2} \cdots \\ a^{m-3}C_{l-2}C_{l-3} \cdots \\ \cdot \quad \cdot \quad \cdot \quad \cdots \\ \cdot \quad \cdot \quad \cdot \quad \cdots \end{vmatrix}. \tag{5}$$

Определитель $\Delta_l(a)$ имеет порядок $m=l+2/2$ для четных l и $m=l+3/2$ для нечетных l.

Правая граница a_{n+1} применимости формулы $P_n(a)$ равна бесконечности, левая граница a_0 применимости формулы $P_0(a)$ равна нулю.

Из структуры решения следует, что добавление очередного момента в (1) улучшает оценку для $P(x \geqslant a)$ лишь для значений x, превышающих наибольший корень уравнения (5).

2. Рассмотрим, как изменяется структура решения при последовательном добавлении к совокупности моментов очередного момента.

1) Задан только момент нулевого порядка $C_0=1$.

$$P(x \geqslant a)=1 \text{ при } 0 \leqslant a \leqslant \infty.$$

2) Заданы моменты $C_0=1$ и C_1.

$$P(x \geqslant a) = \begin{cases} 1 & \text{при} \quad 0 \leqslant a \leqslant C_1; \\ \dfrac{C_1}{a} & \text{при} \quad C_1 \leqslant a \leqslant \infty. \end{cases}$$

3) Заданы моменты $C_0=1$, C_1, C_2.

$$P(x \geqslant a) = \begin{cases} 1 & \text{при} \quad 0 \leqslant a \leqslant C_1; \\ \dfrac{C_1}{a} & \text{при} \quad C_1 \leqslant a \leqslant \dfrac{C_2}{C_1}; \\ \dfrac{C_2 - C_1^2}{a^2 - 2C_1 a + C_2} & \text{при} \quad \dfrac{C_2}{C_1} \leqslant a \leqslant \infty. \end{cases}$$

4) Заданы моменты $C_0=1$, C_1, C_2, C_3.

$$P(x \geqslant a) = \begin{cases} 1 & \text{при} \quad 0 \leqslant a \leqslant C_1; \\ \dfrac{C_1}{a} & \text{при} \quad C_1 \leqslant a \leqslant \dfrac{C_2}{C_1}; \\ \dfrac{C_2 - C_1^2}{a^2 - 2C_1 a + C_2} & \text{при} \quad \dfrac{C_2}{C_1} \leqslant a \leqslant a_3; \\ \dfrac{C_3 C_1 - C_2^2}{C_1 a^3 - 2C_2 a^2 + C_3 a} & \text{при} \quad a_3 \leqslant a \leqslant \infty, \end{cases}$$

где

$$a_3 = \frac{C_3 - C_2 C_1 + \sqrt{(C_3 - C_2 C_1)^2 - 4\left(C_2 - C_1^2\right)\left(C_3 C_1 - C_2^2\right)}}{2\left(C_3 C_1 - C_2^2\right)}$$

и т. д.

3. Для определения по совокупности моментов вероятности $P(|x| \geqslant a)$ можно пользоваться формулами п. 1 раздела, положив

$$C_k = \int_{-\infty}^{\infty} f(x)\, |x|^k \, dx \qquad (k=0, \ldots, n). \tag{6}$$

Если заданы моменты

$$C_k = \int_{-\infty}^{\infty} f(x)\, x^k \, dx \qquad (k=0, \ldots, n), \tag{7}$$

то нечетные моменты не влияют на оценку вероятности и из совокупности моментов следует выбрать только четные. Поэтому задача сводится к рассмотренной в п. 1 для случайной величины x^2. В соответствии с этим в соотношениях для вероятности $P_l(a)$

необходимо удвоить степень a и индексы моментов. Границы применимости формул $P_l(a)$ при этом определяются как $\sqrt{a_l}$ и $\sqrt{a_{l+1}}$.

Заметим, что если заданы моменты (7) и $n=2$, то оценка сверху для $P(|x|>a)$ при $C_0=1$ и $C_1=0$ совпадает с неравенством Чебышева.

Поступило в редакцию 16 июля 1974 г.

УДК 681.883.519.2

В. И. БОРЩЕВИЧ

(Кишинев)

ОБ ОДНОЙ СИСТЕМЕ ЧИСЛОВЫХ ХАРАКТЕРИСТИК ОДНОМЕРНЫХ И УСЛОВНЫХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ

Большинство известных в настоящее время статистических методов анализа амплитудных распределений [1] требует большого объема априорной информации об исследуемом процессе. Это связано с тем, что начальные условия (например, уровни квантования) задаются оптимальным в некотором смысле образом только при наличии ряда предварительных сведений о характере распределения вероятностей данного процесса. Поэтому для тех случаев, когда получение таких сведений затруднительно, желательно синтезировать методы, способные адаптироваться к произвольным законам распределения. Это означает, что алгоритмы таких методов должны использовать получаемую в ходе эксперимента информацию для автоматической установки уровней, оптимальных с той или иной точки зрения. Кроме того, очевидно, что значения этих уровней уже сами по себе будут характеризовать исследуемое распределение.

Рассмотрим один из возможных путей решения поставленной задачи. Пусть имеется априорно неизвестный закон распределения $F(x)$, заданный на интервале исследуемых значений аргумента $[a, b]$. Сначала определяем среднее значение μ_1 (для простоты вместо оценок пока будем говорить о точных значениях), на втором этапе разбиваем интервал $[a, b]$ на два непересекающихся интервала $[a, \mu_1]$ и $[\mu_1 b]$, на которых определяем средние значения μ_{10}, μ_{11} (индексы записываются в двоичной системе счисления), на третьем этапе, используя предыдущие числовые значения, разбиваем $[a, b]$ на интервалы $[a, \mu_{10})$, $[\mu_{10}, \mu_1)$, $[\mu_1, \mu_{11})$ и $[\mu_{11}, b]$, где соответственно находим μ_{100}, μ_{101}, μ_{110}, μ_{111} и т. д.

Назовем совокупности чисел, получаемых на каждом этапе, группами и отметим, что количество цифр в двоичном индексе i соответствует порядковому номеру группы, к которой принадлежит данное число μ_i, а количество таких чисел в m-й группе равно 2^{m-1}. Аналитически через интеграл Стильтьеса

$$\mu_i = \frac{1}{P_i} \int\limits_{\alpha_i}^{\beta_i} x dF(x), \tag{1}$$

где P_i — вероятность $P\{x \in [\alpha_i, \beta_i)\}$; $[\alpha_i, \beta_i)$ — интервал, на котором определяется данное число μ_i.

Оценки $\{\mu_i^*\}$ для сигналов, представленных в дискретной форме, удобно получить с помощью рекуррентного алгоритма вида

$$\mu_i[n] = \mu_i[n-1] + \gamma[n](x[n] - \mu_i[n-1]). \tag{2}$$

Здесь $x[n]$ — n-й по счету сигнал, поступающий на вход i-го канала при условии, что

$$\alpha_i^* \leqslant x[n] < \beta_i^*, \tag{3}$$

где α_i^*, β_i^* — текущие оценки нижнего и верхнего порогов дискриминации. Такой алгоритм [2] обеспечивает сходимость $[\mu_i^*]$ к $[\mu_i]$ в среднеквадратичном, если

$$\sum_{n=1}^{\infty} \gamma[n] = \infty; \quad \sum_{n=1}^{\infty} \gamma^2[n] < \infty. \tag{4}$$

Блок-схема соответствующего устройства представлена на рис. 1, где $Д_i$ — двухпороговый дискриминатор, μ_i^* — блок рекуррентного оценивания.