УДК 519.25

Б.Ю.Лемешко
(Новосиб.электротехн.ин-т)

К ВОПРОСУ РЕШЕНИЯ ЗАДАЧИ АСИМПТОТИЧЕСКИ ОПТИМАЛЬНОГО ГРУППИРОВАНИЯ ДАННЫХ ПРИ ОБРАБОТКЕ НАБЛЮДЕНИЙ, ПОДЧИНЯЮЩИХСЯ БЕТА-РАСПРЕДЕЛЕНИЮ

Статистическая обработка наблюдений часто сопровождается группированием данных, что вызывается либо необходимостью сжатия больших объемов информации, либо используемой процедурой обработки. Известно, что качество получаемых оценок неизвестных параметров и статистических выводов при проверке гипотез существенно зависят от способа группирования, от того, как разбита область определения случайной величины на непересекающиеся интервалы группирования.

Оценки параметров распределений по группированным данным находят обычно методом максимального правдоподобия, методом минимума χ^2 или одним из родственных ему методов [1,2]. При определенных условиях [1,3-6] эти оценки существуют, единственны и обладают свойствами состоятельности и асимптотической эффективности. В этом случае асимптотическая дисперсионная матрица оценки вектора параметров θ распределения определяется соотношением

$$D[\theta] = N^{-1} J_r^{-1}(\theta),$$

где N - объем выборки; $J_r(\theta)$ - информационная матрица Фишера по группированным наблюдениям. Очевидно, что с ростом потерь информации, связанных с группировкой данных, ухудшается и дисперсионная матрица оценок параметров.

В то же время, как показано в работе [7], при проверке гипотез по критериям, использующим группировку наблюдений, мощность критериев тем выше, чем ниже потери информации от группирования.

Матрица $J_r(\theta)$ зависит от граничных точек интервалов группирования x_i, разбивающих область определения случайной величины $\mathcal{X}$ на непересекающиеся интервалы, т.е.

$$\inf \mathcal{X} = x_0 < x_1 < x_2 < \ldots < x_k = \sup \mathcal{X}.$$

Таким образом, задаваясь определенным критерием, можно минимизировать потери информации от группирования, разбивая соответствующим образом область определения случайной величины, т.е. выбирая $x_1, x_2, \ldots, x_{k-1}$.

В данном случае рассматривается решение задачи асимптотически оптимального группирования для одного из наиболее часто встречающихся в приложениях бета-распределения.

Функция плотности бета-распределения задается выражением

$$f(x;\alpha,\beta)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1},$$

где $\alpha,\beta>0$, $x\in[0,1]$, $B(\alpha,\beta)$ – бета-функция, определяемая соотношением

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}=\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt.$$

Информационное количество Фишера о параметрах α распределения по группированным наблюдениям задается соотношением

$$\mathcal{J}_r(\alpha)=\sum_{i=1}^{k}\left(\frac{\partial P_i(\alpha,\beta)}{\partial\alpha}\right)^2/P_i(\alpha,\beta),$$

где вероятность попадания наблюдения в интервал

$$P_i(\alpha,\beta)=\frac{1}{B(\alpha,\beta)}\int_{x_{i-1}}^{x_i}t^{\alpha-1}(1-t)^{\beta-1}dt.$$

После соответствующих преобразований выражение для информационного количества по группированным данным представляется в виде

$$\mathcal{J}_r(\alpha)=\sum_{i=1}^{k}\Big[\frac{1}{B(\alpha,\beta)}\int_{x_{i-1}}^{x_i}\ln t\cdot t^{\alpha-1}(1-t)^{\beta-1}dt-$$

$$-\{\Psi(\alpha)-\Psi(\alpha+\beta)\}P_i(\alpha,\beta)\Big]^2/P_i(\alpha,\beta),$$

где $\Psi(\alpha)$ – логарифмическая производная гамма-функции:

$$\Psi(\alpha)=\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}.$$

Для информационного количества Фишера по негруппированным наблюдениям о параметре α на одно наблюдения имеем

$$\mathcal{J}_H(\alpha)=E\left[\left\{\frac{\partial\ln f(x,\alpha,\beta)}{\partial\alpha}\right\}^2\right]=\Psi'(\alpha)-\Psi'(\alpha+\beta).$$

Задача асимптотически оптимального группирования при оценивании параметра α при известном β или при проверке гипотез о параметре α формулируется следующим образом. Необходимо максимизировать относительную асимпотическую информацию о параметре α :

$$A_1 = J_r(\alpha) / J_H(\alpha)$$

при $0 < x_1 < x_2 < \ldots < x_{k-1} < 1$.

Информационное количество Фишера о параметре β по группированным наблюдениям определяется выражением

$$J_r(\beta) = \sum_{i=1}^{k} \left(\frac{\partial P_i(\alpha,\beta)}{\partial \beta} \right)^2 / P_i(\alpha,\beta) =$$

$$= \sum_{i=1}^{k} \left[\frac{1}{B(\alpha,\beta)} \int_{x_{i-1}}^{x_i} \ln(1-t)\, t^{\alpha-1} (1-t)^{\beta-1} dt - \right.$$

$$\left. -(\Psi(\beta) - \Psi(\alpha+\beta)) P_i(\alpha,\beta) \right]^2 / P_i(\alpha,\beta) .$$

По негруппированным данным соответствующее информационное количество имеет вид

$$J_H(\beta) = E\left[\left\{ \frac{\partial \ln f(x,\alpha,\beta)}{\partial \beta} \right\}^2 \right] = \Psi'(\beta) - \Psi'(\alpha+\beta) .$$

Задача асимптотически оптимального группирования при оценивании параметра β или при проверке гипотез о нем при известном α принимает вид. Найти

$$\max_{0<x_1<x_2<\ldots<x_{k-1}<1} A_2 = \max_{0<x_1<x_2<\ldots<x_{k-1}<1} J_r(\beta) / J_H(\beta) .$$

При оценивании одновременно параметров α и β или при проверке гипотез о согласии рассматривается информационная матрица по группированным наблюдениям

$$J_r = \begin{bmatrix} J_r(\alpha) & J_r(\alpha,\beta) \\ J_r(\alpha,\beta) & J_r(\beta) \end{bmatrix} ,$$

где

$$J_r(\alpha,\beta) = \sum_{i=1}^{k} \left[\frac{\partial P_i(\alpha,\beta)}{\partial \alpha} \quad \frac{\partial P_i(\alpha,\beta)}{\partial \beta} \right] / P_i(\alpha,\beta) =$$

$$= \sum_{i=1}^{k} \frac{1}{P_i(\alpha,\beta)} \left[\frac{1}{B(\alpha,\beta)} \int_{x_{i-1}}^{x_i} \ln t \cdot t^{\alpha-1} (1-t)^{\beta-1} dt - \right.$$

$$\left. -(\Psi(\alpha) - \Psi(\alpha+\beta)) P_i(\alpha,\beta) \right] \times \left[\frac{1}{B(\alpha,\beta)} \int_{x_{i-1}}^{x_i} \ln(1-t) \cdot t^{\alpha-1} (1-t)^{\beta-1} dt - \right.$$

$$\left. -(\Psi(\beta) - \Psi(\alpha+\beta)) P_i(\alpha,\beta) \right] .$$

Соответствующая информационная матрица по негруппированным наблюдениям имеет вид

$$\mathcal{J}_H = \begin{bmatrix} \mathcal{J}_H(\alpha) & \mathcal{J}_H(\alpha,\beta) \\ \mathcal{J}_H(\alpha,\beta) & \mathcal{J}_H(\beta) \end{bmatrix};$$

где

$$\mathcal{J}_H(\alpha,\beta) = E\left[\frac{\partial \ln f(x;\alpha,\beta)}{\partial\alpha}\,\frac{\partial \ln f(x;\alpha,\beta)}{\partial\beta}\right] = -\psi'(\alpha+\beta).$$

Задача асимптотически оптимального группирования в этом случае может быть сформулирована следующим образом. Найти

$$\max_{0<x_1<x_2<\ldots<x_{k-1}<1} A_3 = \max_{0<x_1<x_2<\ldots<x_{k-1}<1} \det \mathcal{J}_r / \det \mathcal{J}_H .$$

причем

$$\det \mathcal{J}_H = \psi'(\alpha)\psi'(\beta) = \psi'(\alpha+\beta)[\psi'(\beta)-\psi'(\alpha)].$$

Как видно из приведенных выше соотношений граничные точки интервалов группирования нельзя представить в виде, инвариантном относительно параметров α и β. Поэтому соответствующие решения задач асимптотически оптимального группирования могут быть получены либо в виде множества таблиц с граничными точками при различных значениях числа интервалов k и различных значениях α и β, либо с использованием разработанного программного обеспечения задач асимптотически оптимального группирования при оценивании или при проверке гипотез.

Литература

1. Рао С.Р. Линейные статистические методы и их применения. - М.: Наука, 1968. - 548 с.

2. Кендали М., Стьюарт А. Статистические выводы и связи. - М.: Наука, 1973. - 900 с.

3. Куллдорф Г. Введение в теорию оценивания по группированным и частично группированным выборкам. - М.: Наука, 1966. - 176 с.

4. Бодин Н.А. Оценка параметров распределений по группированным наблюдениям // - Труды Математического ин-та им. В.А.Стеклова. - Л., 1970. - Т.111. - С.110-154.

5. Лемешко Б.Ю. Оценивание параметров распределений по группированным наблюдениям // Вопросы кибернетики. - М.: 1977. - Вып. 30. - С.80-96.

6. Денисов В.И. Математическое обеспечение системы ЭВМ. - экспериментатор. - М.: Наука, 1977. - 252 с.

7. Денисов В.И., Лемешко Б.Ю. Оптимальное группирование при обработке экспериментальных данных // Измерительные информационные системы / Новосиб. электротехн. ин-т. - Новосибирск, 1979. - С. 5-14.

УДК 621.313.322.81

А.И.Горнинг, М.Н.Задорожная,
Б.А.Иткин
(Новосибирск.электротехн.ин-т)

МОДЕЛИ ВНУТРЕННЕЙ ПРОПИТКИ ИЗОЛЯЦИИ ОБМОТКИ СТАТОРА ТУРБО-И ГИДРОГЕНЕРАТОРОВ

При изготовлении турбо-и гидрогенераторов важно и в то же время наиболее трудно получить высокое качество изоляции обмотки статора. Особенно остро этот вопрос стоит для высоковольтных генераторов (с большими толщинами изоляции), изготовляемых по технологии вакуум-нагнетательной пропитки, при которой в ряде случаев имеет место недопропитка внутренних слоев изоляции. С целью устранения этого недостатка нами был предложен ряд специальных конструктивно-технологических изменений, позволяющих более эффективно использовать процесс внутренней пропитки изоляции стержней обмотки статора [1]. В связи с этим, наряду с предложенными ранее моделями пропитки [2], необходимо иметь модели только внутренней пропитки, позволяющие на стадии проектирования генераторов и создания технологии принять оптимальные решения для получения высокого качества обмоток статоров.

Процесс внутренней пропитки изоляции стержня осуществляется следующим образом. Первоначально компаунд поступает в каналы 1, образующиеся между элементарными проводниками, и специальный канал 2, полученный путем удаления одного из элементарных проводников (рис. 1,2). Далее пропитывающий состав заполняет каналы 3, образующиеся на стыках лент (рис.2). Происходит последовательная пропитка слоев изолирующей ленты. Таким образом, процесс внутренней пропитки можно принять состоящим из двух основных стадий: движения вязкой несжимаемой жидкости в канале и проникновения компаунда в пористую среду (изоляцию).