У цьому прикладі оптимальні k та l виявилися різними для різних підінтервалів. Це дивно, бо звичайно для $\Gamma(x)$ використовується раціональне наближення при $k=l$ [5]. Проте наближення тієї ж функції на тих же підінтервалах раціональним виразом $R_{1,1}(x)$ дає такі максимальні похибки $\mu_1 = .015051$, $\mu_2 = .0033367$, $\mu_3 = .0013161$. Результати виявляються гіршими на всіх підінтервалах. Найбільша різниця – у 12 разів.

1. Попов Б.А. Равномерное приближение сплайнами. Киев: Наук.думка, 1989. 272с. 2. Попов Б.О., Лаушник О.І. Знаходження оптимальної форми раціональних наближень// Відбір і обробка інформ. 1997. Вип.11(87). С.107-110. 3. Meinardus G. Some Results in Segmented Approximation// Comput. Math. Appl. 1997. Vol.33. №1/2. P.165-180. 4. Monagan M.B., Geddes K.O., Heal K.M. et al. Maple V. Programming Guide. N.Y.: Springer-Verlag. 1998. 390p. 5. Moshier S.L. Methods and Programs for Mathematical Functions. N.Y.: John Wiley. 1989. 418p. 6. Nurnberger G. Approximations by Spline Functions. N.Y.: Springer-Verlag. 1989. 416p. 7. Popov B.A. A System for Function Approximation Properties Investigation //SIGSAM Bulletin. 1997. Vol.31. №3. P.29-30.

УДК 519.6:518

П.С. Сеньо

РОЗВ'ЯЗАННЯ ОСНОВНОЇ ЗАДАЧІ МАТЕМАТИЧНОЇ СТАТИСТИКИ МЕТОДАМИ ІНТЕРВАЛЬНОГО АНАЛІЗУ

Аналіз статистичного матеріалу закінчено, коли істотні властивості генеральної сукупності записано аналітично. Зокрема, ця основна задача математичної статистики розв`язана повністю, коли відома функція розподілу, або густина генеральної сукупності. Практично, ці функції дослідник *підбирає* на основі аналізу вибірки - відомі критерії погодженості (Пірсона, Колмогорова, Смірнова тощо) лише з певним ступенем надійності обгрунтовують зроблений вибір; можемо лише *оцінити* значення невідомих параметрів розподілу та й це лише при умові, коли сам розподіл відомий.

Суть проблеми полягає в слідуючому. Потрібно побудувати таку криву, графік якої "найщільніше" пройшов би через множину точок (x_i, n_i); побудована крива повинна бути *густиною*; крива повинна бути теоретично обгрунтованою; найбільш простою; рівняння кривої повинне містити найменше можливе число параметрів. Тут x_i - елементи частотної таблиці вибірки; n_i - вибіркові частоти. Така крива обгрунтовано вважається розподілом генеральної сукупності і може використовуватися як в практиці (наприклад, можна обчислити n_i і для x_i, яких і немає в таблиці; розв`язувами задачі інтерполяції, екстраполяції тощо) так і при теоретичних дослідженнях.

Криву, яка вирівнює точки (x_i, y_i) площини можна побудувати нескінченною кількістю способів (це різні криві !). Тому підбираємо таку криву, але лише з певної фіксованої множини кривих. Виходячи з теореми Вейєрштраса, найчастіше підбирають параболу $y = a_0 + a_1 x + \cdots + a_m x^m$. Однак цей шлях нам не підходить в зв`язку з тим, що для k точок така крива має степінь $k-1$, а в нас, як правило, k велике ($\approx$ 10 - 1000); така парабола, як правило, не є густиною; похибки такої апроксимації великі.

Більшість розподілів випадкових змінних характеризуються такими властивостями:

1. значення таких змінних лежать в певних межах, тобто, ймовірності $P(\zeta < \zeta_{\min}) = P(\zeta > \zeta_{\max}) = 0$;

2. густина розподілу є унімодальною функцією - на проміжку $[\zeta_{\min}, a]$ вона зростає, в точці a досягає максимуму і далі на проміжку $[a, \zeta_{\max}]$ спадає;

3. на кінцях проміжка $[\zeta_{\min}, \zeta_{\max}]$ похідна від густини рівна нулю та й часто дотик графіка густини $p(x)$ до осі OX є вищого порядку.

К.Пірсон показав [3], що більшість розподілів ймовірностей є розв`язками такого диференціального рівняння

$$y' = y(x+a)/(b_0 + b_1 x + b_2 x^2). \tag{1}$$

До таких розподілів, зокрема, належать такі фундаментальні розподіли як рівномірний, експонентний, нормальний, гама розподіл, бета розподіли 1-го і другого роду; розподіли Пірсона, Стьюдента, Фішера, Парето. Сюди ж належить і будь-який з розподілів, що отримується лінійним перетворенням випадкової змінної, що керується одним з перерахованих розподілів. Однак, практичне застосування цієї методики наштовхується на проблему визначення теоретичних моментів розподілу через статистичні моменти. Цю проблему пом`якшив Л.М.Большов за допомогою асимптотичних перетворень [1].

Нехай функція $y = p(x)$ є розв`язком рівняння

$$\frac{dy}{dx} = y\,\frac{(x+a)}{f(x)}, \tag{2}$$

де $f(x)=\sum_{i=0}^{\infty} b_i x^i$ - деяка функція. Тоді розв`язок рівняння (2) є густиною, що задовольняє умови 1) - 3), якщо коефіцієнти розкладу $\{b_i\}_{i=0}^{\infty}$ виражаються через теоретичні моменти розподілу $m_i = \int_{-\infty}^{\infty} x^i p(x)dx$ так:

$$a\, m_n + n\, b_0 m_{n-1} + (n+1) b_1\, m_n + (n+2) b_2 m_{n+1} + \cdots = -m_{n+1}, \tag{3}$$

де ($n = 0,1,2,\ldots$).

Якщо i-та похідна рівна нулю, або має двохстороннє обмеження, то тоді можна описати всі класи розподілів, які є узагальненнями кривих Пірсона. Зокрема, для цього достатньо, щоб існувало інтервальне розширення другої похідної функції $f(x)$. Очевидно, що інтервали - розв`язки природно звужуються з врахуванням властивостей густини. Моменти цих розподілів задовольняють наступну систему рівнянь:

$$\begin{cases} am_0+0f(0)+f'(0)m_0+2\frac{1}{2!}[f''(x)]m_1=-m_1, \\ am_1+f(0)m_0+2f'(0)m_1+3\frac{1}{2!}[f''(x)m_2=-m_2, \\ am_2+2f(0)m_1+3f'(0)m_2+4\frac{1}{2!}[f''(x)m_3=-m_3, \\ am_3+3f(0)m_2+4f'(0)m_3+5\frac{1}{2!}f''(x)]m_4=-m_4. \end{cases} \tag{4}$$

Це система рівнянь з інтервальними коефіцієнтами. Розв`язавши її, отримуємо інтервали для коефіцієнтів і межі, коли корені дійсні, кратні, комплексні. Так отримуємо інтервальні аналоги розподілів, які найкраще вирівнюють наявні вибіркові значення. Можливі випадки, коли отримані інтервали перетинаються. Це означає, що в зв`язку з різними збуреннями, вибіркові значення можна вирівнювати обома відповідними розподілами (вони, як правило, асимптотично еквівалентні).

Моменти в системі (4) теоретичні і тому нам невідомі.З вибірки ми можемо отримати всі потрібні статистичні моменти. Інтеграли, якими є теоретичні моменти, оцінюємо інтервально за типовою методикою. Розбивши проміжок інтегрування в цих інтегралах вибірковими значеннями з частотної таблиці отримуємо інтервальні оцінки теоретичних моментів через статистичні. В результаті в системі (4) і величини $m_i (i = 0,1,2,3,4)$ також є відомими інтервалами.

Частотна таблиця *групує* експериментальні дані з неперервної генеральної сукупності. Моменти цього "групового" розподілу відрізняються від моментів неперервного розподілу [2]. Інтервальний варіант поправок Шеппарда отримуємо реалізуючи інтервальну оцінку правої частини рівності

$$\int_{y_i-h/2}^{y_i+h/2} x^k p(x)dx = h(y_i^k p(y_i)) + h^3 \frac{1}{2^2 2!}(y_i^k p(y_i))' + \ldots +$$

$$+ h^{2j-1} \frac{1}{2^{2j-2}(2j-2)!}\left(y_i^k p(\theta_i)\right)^{(2j-2)},$$

де y_i – середини інтервалів таблиці; h – крок таблиці.

1. Большев Л.Н., Асимтотически пирсоновские преобразования, Теория векроятностей и ее применения, том УIII, вып. 2, 1963, с.129-155. 2. *Крамер Г.*, Математические методы статистики., Мир., М., 1975. 3. Романовский В., Математическая статистика., ГОНТИ НКТП СССР, 1938.

УДК 519.689

Д.О. Тарасов

ЗАСОБИ ЗАБЕЗПЕЧЕННЯ ЦІЛІСНОСТІ ДАНИХ НА ОСНОВІ ПРОМИСЛОВИХ СТАНДАРТІВ

Різнопланові інформаційні системи на базі реляційних СУБД широко використовуються для автоматизації обліку, опрацювання та аналізу інформації. Метою роботи таких систем є спрощення процесів прийняття рішень за рахунок використання інформації з БД. Однак рішення має грунтуватися на несуперечливій інформації. Таким чином, для забезпечення результативності дії інформаційної системи СУБД повинна забезпечувати цілісність даних у БД. Різні СУБД використовували власні засоби для забезпечення цілісності даних. Їх засоби розвивались відповідно до специфіки застосування, забезпечення сумісності версій та політики фірми виробника.