УДК 519.21

ДОСЛІДЖЕННЯ ОЦІНОК ДЕЯКИХ ХАРАКТЕРИСТИК ВИПАДКОВИХ ВЕЛИЧИН

Тупко Н.П. *(Одеська державна академія будівництва та архітектури, м.Одеса)*

У роботі досліджені зміщені і незміщені оцінки змішаного моменту другого порядку та коефіцієнта коваріації. Отримані умови, за яких одна з оцінок краща, ніж інша для нормально розподілених випадкових величин.

Для розв`язку різноманітних технологічних задач будівництва дуже часто використовуються ймовірносно-статистичні методи. Досить успішно ці методи розвиваються у двох основних напрямках:

- статистичний аналіз накопичених даних про технологічні процеси, рецептури та властивості сировини і матеріалів з ціллю узагальнення інформації у вигляді формул;
- розробка математично-статистичних моделей для прийняття оптимальних технологічних та економічних рішень, а також для управління процесом чи окремими його етапами, зокрема з використанням ЕОМ.

І слід зазначити, що при вирішенні того чи іншого питання виникає потреба оцінити основні характеристики випадкових величин. До таких характеристик відносять змішаний момент другого порядку та коефіцієнт коваріації. Наприклад, хочемо визначити коефіцієнт кореляції (який визначає силу зв'язку між випадковими величинами) між границею тривкості на стиснення (x) та теплостійкістю (y) деякого матеріалу по експериментальним даним, то для цього треба використати оцінку коефіцієнта коваріації і чим краща буде оцінка, тим точніший результат.

Нехай x, y пара випадкових величин, $K(x,y)=m[x-m(x)][y-m(y)]$ - коефіцієнт коваріації, де $m(x)$ та $m(y)$- математичне сподівання відповідно x та y. За означенням $K(x,y)=m(xy)-m(x)m(y)$, тобто для оцінки коефіцієнта коваріації $K(x,y)$ необхідно оцінити змішаний момент другого порядку $\gamma=m(xy)$ та математичне сподівання $m(x), m(y)$.

Розглянемо вибірки пар $(x_1, y_1),\ldots,(x_n y_n)$ - послідовність значень незалежних однаково розподілених випадкових величин.

Проблему оцінки невідомого коефіцієнта коваріації можна досліджувати у випадку відомих математичних сподівань $M(x) = \alpha, M(y) = \beta$, якщо об`єм n-вибірки більше тридцяти, а також у випадку невідомих математичних сподівань, якщо об`єм менше тридцяти.

Розглянемо клас білінійних оцінок змішаного моменту другого порядку $\gamma = M(xy)$:

$$H = \left\{ u = \sum_{i,k=1}^{n} a_{ik} x_i y_k = (Ax, y), A = \|a_{ik}\|_{i,k=1,n}, x = (x_{1,\ldots,}x_n), y = (y_1,\ldots,y_n) \right\}$$

і дві оцінки з цього класу: $\gamma(x,y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$ - незміщену та

$\tilde{\gamma}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n} x_i y_i$ - зміщену.

Оцінка $\gamma(x,y)$- незміщена, так як

$$m(\gamma(x,y)) = m\left[\frac{1}{n}\sum_{k=1}^{n} x_k y_k\right] = \frac{1}{n} m\left[\sum_{k=1}^{n} x_k y_k\right] = \frac{1}{n}\cdot n\gamma(x,y) = \gamma(x,y)$$

Оцінка

$$m(\tilde{\gamma}(x,y)) = m\left(\frac{n}{n-1}\gamma(x,y)\right) = \frac{n}{n-1}\gamma(x,y) = \frac{n}{n-1}\gamma(x,y) \quad \text{зміщена.}$$

Порівняємо середньоквадратичне відхилення цих оцінок, тобто з`ясуємо яка оцінка і при яких умовах краща.

$$V = \delta^2(\gamma) - \delta^2(\tilde{\gamma}) = m\left[(\gamma-\gamma)^2\right] - m\left[(\tilde{\gamma}-\gamma)^2\right] = m\left[(\gamma-\gamma)^2 - (\tilde{\gamma}-\gamma)^2\right] =$$
$$= m[\ (\gamma-\gamma-\tilde{\gamma}+\gamma)(\gamma-\gamma+\tilde{\gamma}-\gamma)\ ] = m[\ (\gamma-\tilde{\gamma})(\gamma+\tilde{\gamma}-2\gamma)\ ] =$$
$$= m\left[\ \left(\gamma - \frac{n}{n-1}\tilde{\gamma}\right)\left(\gamma + \frac{n}{n-1}\tilde{\gamma} - 2\gamma\right)\ \right] = -\frac{1}{n-1} m\left[\ \gamma\left(\frac{2n-1}{n-1}\gamma - 2\gamma\right)\ \right] =$$

$$= -\frac{1}{n-1} m\left[\ \frac{2n-1}{n-1}\gamma^2 - 2\gamma\gamma\ \right] = -\frac{2n-1}{n-1} m(\gamma^2) + \frac{2}{n-1}\gamma^2 .$$

Обчислимо $m(\gamma^2)$:

$$m\left(\gamma^2(x,y)\right)=m\left(\frac{1}{n^2}\left(\sum_{i=1}^{n}x_iy_i\right)^2\right)=\frac{1}{n^2}m\left(\sum_{i=1}^{n}x_i^2y_i^2+\sum_{\substack{i,k=1\\i\neq k}}^{n}x_iy_ix_ky_k\right)=$$

$$=\frac{1}{n^2}\left(m\left(\sum_{i=1}^{n}x_i^2y_i^2\right)+m\left(\sum_{\substack{i,k=1\\i\neq k}}^{n}x_iy_ix_ky_k\right)\right)=\frac{1}{n^2}\Big(n(n-1)m^2(xy)+$$

$$+nm\left(x^2y^2\right)\Big)=\frac{n-1}{n}\gamma^2+\frac{1}{n}m\left(x^2y^2\right)$$

Розглянемо двовимірний нормальний розподіл з $m(x)=0, m(y)=0$ і $D(x)=1, D(y)=1$ та ρ-коефіцієнтом кореляції. Відповідно змішані моменти мають наступний вид:

$$m(xy)=\rho\,,\; m(xy^2)=m(x^2y)=0\,,\; m(x^2y^2)=1-2\rho^2\,. \tag{1}$$

Тоді

$$m\left(\gamma^2(x,y)\right)=\frac{n-1}{n}\rho^2+\frac{1}{n}\left(1-2\rho^2\right)=\frac{n-1}{n}\rho^2+\frac{1}{n}-\frac{2\rho^2}{n}=\frac{n-3}{n}\rho^2+\frac{1}{n}$$

В цьому випадку

$$V=-\frac{2n-1}{n-1}\left(\frac{n-3}{n}\rho^2+\frac{1}{n}\right)+\frac{2}{n-1}\rho^2=-\frac{1}{n-1}\left(\frac{2n-1}{n-1}\left(\frac{(n-3)}{n}\rho^2+\right.\right.$$

$$\left.\left.+\frac{1}{n}\right)-2\rho^2\right)=-\frac{1}{n-1}\left(\frac{-5n\rho^2+3\rho^2+2n-1}{n(n-1)}\right)=$$

$$=-\frac{1}{n(n-1)^2}\left[n\left(2-5\rho^2\right)+3\rho^2-1\right]$$

При умові $n\geq 2$ і $\rho^2\leq\frac{2}{5}$, різниця - від'ємна $V<0$, тобто незміщена оцінка $\gamma(x,y)$ краща, ніж зміщена $\bar{\gamma}(x,y)$. А при умові $n\geq 2$ і $\rho^2>\frac{2}{5}$, зміщена оцінка $\bar{\gamma}(x,y)$ буде краще незміщеної $\gamma(x,y)$.

Тепер розглянемо дві оцінки коефіцієнта коваріації незміщену та зміщену:

$$K(x,y)=\frac{1}{n}\sum_{i=1}^{n}(x_i-\alpha)(y_i-\beta)$$

та

$$\tilde{K}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\alpha)(y_i-\beta)\,.$$

Оцінка $K(x,y)$ - незміщена, так як

$$m(K(x,y)) = m\left[\frac{1}{n}\sum_{k=1}^{n}(x_k-\alpha)(y_k-\beta)\right] = \frac{1}{n}m\left[\sum_{k=1}^{n}(x_k y_k-\alpha y_k-\beta x_k+\alpha\beta)\right] =$$

$$= \frac{1}{n}\left[nm(xy)-\alpha nm(y)-\beta nm(x)+\alpha\beta n\right] = \gamma-\alpha\beta-\beta\alpha+\alpha\beta = \gamma-\alpha\beta = K(xy).$$

Оцінка

$$m(\tilde{K}(x,y)) = m\left(\frac{n}{n-1}K(x,y)\right) = \frac{n}{n-1}K(x,y) \quad \text{зміщена.}$$

Для того щоб дослідити точність оцінок $K(x,y)$ і $\tilde{K}(x,y)$, необхідно порівняти середньоквадратичне відхилення цих оцінок. З цією ціллю розглянемо співвідношення:

$$V_1 = \delta^2(K)-\delta^2(\tilde{K}) = m\left[(K-K)^2\right]-m\left[(\tilde{K}-K)^2\right] = m\left[(K-K)^2-(\tilde{K}-K)^2\right] =$$

$$= m\left[\ (K-K-\tilde{K}+K)(K-K+\tilde{K}-K)\ \right] = m\left[\ (K-\tilde{K})(K+\tilde{K}-2K)\ \right] =$$

$$= m\left[\ \left(K-\frac{n}{n-1}K\right)\left(K+\frac{n}{n-1}K-2K\right)\ \right] = -\frac{1}{n-1}m\left[\ K\left(\frac{2n-1}{n-1}K-2K\right)\ \right] =$$

$$= -\frac{1}{n-1}\left[\ \frac{2n-1}{n-1}m\left(K^2\right)-2K^2\ \right].$$

Підрахуємо $m\left(K^2\right)$:

$$m\left(K^2(x,y)\right) = m\left(\frac{1}{n^2}\left(\ \sum_{i=1}^{n}(x_i-\alpha)(y_i-\beta)\ \right)^2\right) = \frac{1}{n^2}m\left(\ \sum_{i=1}^{n}(x_i-\alpha)^2(y_i-\beta)^2 +\right.$$

$$\left.+\ \sum_{\substack{i,j=1\\ i\neq j}}^{n}(x_i-\alpha)(y_i-\beta)(x_k-\alpha)(y_k-\beta)\ \right) = \frac{1}{n^2}m\left(\ \sum_{\substack{i,j=1\\ i\neq j}}^{n}(x_i y_i x_k y_k-\alpha x_i y_i y_k +\right.$$

$$-\beta x_i y_i x_k+\alpha\beta x_i y_i-\alpha x_k y_k y_i+\alpha^2 y_i y_k-\alpha^2\beta y_i-\beta x_i x_k y_k+\alpha\beta x_i y_k-\beta^2 x_i x_k -$$

$$-\beta^2\alpha x_i + \alpha\beta x_k y_k - \alpha^2\beta y_k - \alpha\beta^2 x_k + \alpha^2\beta^2 \ ) + \sum_{i=1}^{n}\left(x_i^2 y_i^2 - 2\beta x_i^2 y_i + \beta^2 x_i^2 - \right.$$

$$\left. -2\alpha x_i y_i^2 + 4\alpha\beta x_i y_i - 2\alpha\beta^2 x_i + \alpha^2 y_i^2 - 2\alpha\beta^2 y_i + \alpha^2\beta^2 \ ) \right) =$$

$$= \frac{1}{n^2}\left(n(n-1)\left(m^2(xy) + 2\alpha^2\beta^2 + + \alpha\beta m(xy)\right) + n\left(m\left(x^2y^2\right) - 2\beta m\left(x^2 y\right) - \right.\right.$$

$$\left.\left. -2\alpha m\left(xy^2\right) + 4\alpha\beta m(xy) + \beta^2 m\left(x^2\right) + \alpha^2 m\left(y^2\right) - 3\alpha^2\beta^2 \right) \right).$$

Розглянемо двовимірний нормальний розподіл (1), тоді

$$V_1 = -\frac{1}{n-1}\left[\frac{2n-1}{n-1}\left(\frac{n-1}{n}\rho^2 + \frac{1}{n}\left(1-2\rho^2\right) \right) - 2\rho^2 \right] =$$

$$= -\frac{1}{n(n-1)}\left[n\left(-5\rho^2+2\right)+3\rho^2-1\right]$$

При умові $n \geq 2$ і $\rho^2 \leq \frac{2}{5}$, різниця - від'ємна $L < 0$, тобто незміщена оцінка $K(x,y)$ краща, ніж зміщена $\tilde{K}(x,y)$. А при умові $n \geq 2$ і $\rho^2 > \frac{2}{5}$, зміщена оцінка $\tilde{K}(x,y)$ буде краще незміщеної $\dot{K}(x,y)$.

Висновок. Таким чином отримані умови при яких одна оцінка краще іншої у випадку нормально розподілених випадкових величин. Слід зазначити, що як для змішаного моменту, так і для коефіцієнта коваріації ці умови співпадають, хоча розглянуті оцінки належать різним класам і оцінюють різні характеристики. Для подальшого вивчення цікава поведінка розглянутих оцінок і для інших розподілів.

Література

1. Ю.И.Петунин. Приложение теории случайных процессов в биологии и медицине -К: Наукова думка,1981. -320 с.

2. Петунин Ю.И., Тупко Н.П. Теория квадратичных оценок дисперсии. // Український математичний журнал.-1999.-Том 51 № 9.-С.1217-1231.

3. Ю.Г.Курицын, Ю.И.Петунин. К теории линейных оценок математического ожидания случайного процесса// Теория вероятности и математическая статистика.-1970.-Вып.3.-С.80 - 92.

4. В.Вознесенский. Статистические решения в технологических задачах.-Кишенёв.:Картя Молдовеняске,1969.-232с.

5.Гарольд Крамер. Математические методы статистики.-М.:Мир,1975.-648с.