ПРЕЦИЗНОСТ СТАНДАРДНЕ ДЕВИЈАЦИЈЕ КОД МАКАКВОГ РАСПОРЕДА

БРАНИСЛАВ ИВАНОВИЋ (Београд)

Проф. Кашанин у својој расправи „Le coefficient d'approximat on moyenne et le coefficient de corrélation"[1] напоменуо је на стр. 74 да је израчунавање помоћу стандардне девијације σ_2 у *општем случају* неприцизније него помоћу σ_{2k} за $k > 1$.

Овде ћемо показати да заиста можемо у извесним случајевима добити помоћу стандардне девијације већу прецизност него помоћу σ_{2k} $(k > 1)$.

I

Нека су $a_\nu (\nu = 1, 2, \ldots, n)$ реални бројеви међу којима може бити и једнаких, али нису сви једнаки. Узмимо аритметичку средину $a = \frac{1}{n} \sum_{\nu=1}^{n} a_\nu$ и образујмо разлику

$$\alpha_\nu = a_\nu - a .$$

Ставимо

$$\sigma_{2k} = + \sqrt[2k]{\frac{1}{n} \sum_{\nu=1}^{n} \alpha_\nu^{2k}}$$

где је k цео број. За $k = 1$ имамо стандардну девијацију

$$\sigma_2 = + \sqrt{\frac{1}{n} \sum_{\nu=1}^{n} \alpha_\nu^2}.$$

Како смо претпоставили да нису сви α_ν нуле, то је сигурно $\sigma_{2k} > 0$ за свако k.

Доказаћемо прво ову теорему:

---

[1]) *Publ. de l'Inst. math. de l'Acad. serbe*, t. I, стр. 71—87, 1947.

Теорема I. — *Стандардна девијација је доња граница низа*

$$\sigma_{2k} = \sqrt[2k]{\frac{1}{n}\sum_{\nu=1}^{n} \alpha_\nu^{2k}}, \qquad (k=1,2,\dots),$$

тј.

$$\sigma_2 = \min_{1 \leqslant k < \infty} \sigma_{2k}.$$

Доказ. Имамо

$$\frac{\sigma_{2k}}{\sigma_2} = \frac{\sqrt[2k]{\frac{1}{n}\sum_{\nu=1}^{n}\alpha_\nu^{2k}}}{\sqrt{\frac{1}{n}\sum_{\nu=1}^{n}\alpha_\nu^2}} = \left\{\frac{n^{k-1}\sum_{\nu=1}^{n} a_\nu^{2k}}{\left(\sum_{\nu=1}^{n} a_\nu^2\right)^k}\right\}^{1/2k}.$$

Овај ће израз бити већи од јединице ако је поткорена вредност већа од јединице. За $k=1$ поткорена вредност једнака је јединици; показаћемо да ће за $k=2$ бити већа од јединице. У том циљу писаћемо

$$\frac{n\sum_{\nu=1}^{n}\alpha_\nu^4}{\left(\sum_{\nu=1}^{n}\alpha_\nu^2\right)^2} = \frac{\sum_{\nu=1}^{n}\alpha_\nu^4 + (n-1)\sum_{\nu=1}^{n}\alpha_\nu^4}{\sum_{\nu=1}^{n}\alpha_\nu^4 + 2\sum_{\nu=1,\ \mu=1}^{n}\alpha_\nu^2\alpha_\mu^2}, \qquad (\nu \neq \mu).$$

Како изрази

$$(n-1)\sum_{\nu=1}^{n}\alpha_\nu^4 \quad \text{и} \quad 2\sum_{\nu=1,\ \mu=1}^{n}\alpha_\nu^2\alpha_\mu^2, \qquad (\nu\neq\mu),$$

имају исти број чланова и претстављају симетричне функције и како је увек

$$\alpha_\nu^4 + \alpha_\mu^4 > 2\alpha_\nu^2\alpha_\mu^2$$

кад је $\alpha_\nu \neq \alpha_\mu$, то је

$$(n-1)\sum_{\nu=1}^{n}\alpha_\nu^4 > 2\sum_{\nu=1,\ \mu=1}^{n}\alpha_\nu^2\alpha_\mu^4,$$

те је

$$\frac{n\sum_{\nu=1}^{n}\alpha_\nu^4}{\left(\sum_{\nu=1}^{n}\alpha_\nu^2\right)^2} > 1,$$

тј.

$$\frac{\sigma_4}{\sigma_2} > 1.$$

Ако сада претпоставимо да неједнакост $\sigma_{2k} : \sigma_2 > 1$ вреди за неко k, тј. да је

$$\frac{n^{k-1} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k}}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^k} > 1,$$

доказаћемо да ће та неједначина вредети и за $k+1$, тј. да је и

$$\frac{n^{k} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k+2}}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^{k+1}} > 1. \tag{1}$$

Овај израз можемо написати у облику

$$\frac{n^{k} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k+2}}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^{k+1}} = \frac{n^{k} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k} \alpha_\nu^2}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^{k} \cdot \sum\limits_{\nu=1}^{n} \alpha_\nu^2} =$$

$$= \frac{n^{k-1} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k} \alpha_\nu^2}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^{k} \cdot \frac{1}{n} \sum\limits_{\nu=1}^{n} \alpha_\nu^2} = \frac{n^{k-1} \sum\limits_{\nu=1}^{n} \alpha_\nu^{2k} \frac{\alpha_\nu^2}{\sigma_2^2}}{\left(\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right)^{k}}.$$

Да бисмо, дакле, доказали (1), довољно је показати да се збир $\sum\limits_{\nu=1}^{n} \alpha_\nu^{2k}$ неће смањити ако му сваки члан респективно помножимо са $\alpha_\nu^2 / \sigma_2^2$, тј. показати да је разлика

$$\sum_{\nu=1}^{n} \alpha_\nu^{2k} \frac{\alpha_\nu^2}{\sigma_2^2} - \sum_{\nu=1}^{n} \alpha_\nu^{2k} > 0. \tag{2}$$

Ако из израза на левој страни извучемо $\left\{\sum\limits_{\nu=1}^{n} \alpha_\nu^2\right\}^{-1}$, добићемо

$$\frac{1}{\sum\limits_{\nu=1}^{n} \alpha_\nu^2} \cdot \left\{ n \sum_{\nu=1}^{n} \alpha_\nu^{2k+2} - \sum_{\nu=1}^{n} \alpha_\nu^2 \sum_{\nu=1}^{n} \alpha_\nu^{2k} \right\} =$$

$$= \frac{1}{\sum\limits_{\nu=1}^{n} \alpha_\nu^2} \left\{ n \sum_{\nu=1}^{n} \alpha_\nu^{2k+2} - \sum_{\nu=1}^{n} \alpha_\nu^{2k+2} - \sum_{\nu=1,\, \mu=1}^{n} \alpha_\nu^2 \alpha_\mu^{2k} \right\} =$$

$$= \frac{1}{\sum\limits_{\nu=1}^{n} \alpha_\nu^2} \left\{ (n-1) \sum_{\nu=1}^{n} \alpha_\nu^{2k+2} - \sum_{\nu=1,\, \mu=1}^{n} \alpha_\nu^2 \alpha_\mu^{2k} \right\}.$$

Како су и овде изрази $(n-1)\sum_{\nu=1}^{n}\alpha_\nu^{2k+2}$ и $\sum_{\nu=1,\,\mu=1}^{n}\alpha_\nu^2\alpha_\mu^{2k}$ симетричне функције и имају исти број чланова и како постоје ν и μ тако да је $\alpha_\nu \neq \alpha_\mu$, тј. да је

$$(\alpha_\nu^{2k}-\alpha_\mu^{2k})(\alpha_\nu^2-\alpha_\mu^2)>0,$$

то постоје α_ν и α_μ тако да је

$$\alpha_\nu^{2k+2}+\alpha_\mu^{2k+2}>\alpha_\nu^2\alpha_\mu^{2k}+\alpha_\nu^{2k}\alpha_\mu^2.$$

Зато је

$$\frac{1}{\sum\limits_{\nu=1}^{n}\alpha_\nu^2}\left\{(n-1)\sum_{\nu=1}^{n}\alpha_\nu^{2k+2}-\sum_{\nu=1,\,\mu=1}^{n}\alpha_\nu^2\alpha_\mu^{2k}\right\}>0.$$

Тако смо доказали (2), па тиме и (1).

На тај начин смо путем потпуне индукције доказали теорему I.

Теорема II. — *Ако је*

$$\sigma_{2k}<c<\left(\frac{\sigma_{2k}}{\sigma_2}\right)^{\frac{k}{k-1}}\sigma_2,$$

онда је већи број чланова за које можемо тврдити помоћу стандардне девијације σ_2 да леже у интервалу $(a-c,\ a+c)$ него помоћу σ_{2k} $(k>1)$.

Доказ. — За $k>1$ према теореми I је

$$\left(\frac{\sigma_{2k}}{\sigma_2}\right)^{k-1}<\left(\frac{\sigma_{2k}}{\sigma_2}\right)^{k},\quad \text{тј.}\quad \sigma_{2k}<\left(\frac{\sigma_{2k}}{\sigma_2}\right)^{\frac{k}{k-1}}\sigma_2.$$

Према томе, постоји број c тако да је

$$\sigma_{2k}<c<\left(\frac{\sigma_{2k}}{\sigma_2}\right)^{\frac{k}{k-1}}\sigma_2 \tag{3}$$

Ставићемо $K=\frac{c}{\sigma_2}$, тада је $K>\frac{\sigma_{2k}}{\sigma_2}>1$. Зато у отвореном интервалу $(a-c,\ a+c)$ има бар

$$x=n-\frac{n}{K^2}$$

чланова датог скупа a_ν. Ставимо ли пак $K_1 = \frac{c}{\sigma_{2k}}$ биће

$$K_1 > \frac{\sigma_{2k}}{\sigma_{2k}} = 1$$

па у истом интервалу има бар

$$y = n - \frac{n}{K_1^{2k}} \text{ [1]}$$

чланова датог скупа a_ν. Одатле излази

$$x - y = \frac{n}{K^2 K_1^{2k}} (K^2 - K_1^{2k}).$$

Међутим је

$$K^2 - K_1^{2k} = \frac{c^2}{\sigma_2^2} - \frac{c^{2k}}{\sigma_{2k}^{2k}} = \frac{c^2}{\sigma_2^2 \sigma_{2k}^{2k}} (\sigma_{2k}^{2k} - c^{2k-2}\, \sigma_2^2) >$$

$$> \frac{c^2}{\sigma_2^2 \sigma_{2k}^{2k}} \left\{ \sigma_{2k}^{2k} - \left(\frac{\sigma_{2k}}{\sigma_2}\right)^{\frac{k}{k-1} \cdot 2(k-1)} \sigma_2^{2k-2} {\sigma_2}^2 \right\} = 0,$$

те је $K^2 - K_1^{2k} > 0$, тј. $x > y$. Тиме је тврђење доказано.

На пример за бројеве

$$-5,\ -4,\ -4,\ -2,\ -1,\ -1,\ 1,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3$$

имамо

$$\sigma_2 = 2{,}880,\quad \sigma_4 = 3{,}284,\quad \sigma_6 = 3{,}571,\quad \sigma_8 = 3{,}786,\quad \sigma_{10} = 3{,}952, \ldots$$

Извршићемо компарацију прецизности σ_2 са σ_{10}. Како је

$$\left(\frac{\sigma_{10}}{\sigma_2}\right)^{5/4} = 1{,}485,$$

то треба за c узети неку вредност из интервала (3,952; 4,277). Узмимо, рецимо, $c = 4{,}1$; тада је $K = 1{,}4236$ и $K_1 = 1{,}037$. Помоћу σ_2 можемо тврдити да се у интервалу $(-4{,}1;\ 4{,}1)$ налази чланова бар

$$x = n - \frac{n}{K^2} = 13 - \frac{13}{1{,}4236^2} = 6{,}49,$$

док помоћу σ_{10} можемо утврдити да се у том истом интервалу налази чланова бар

$$y = n - \frac{n}{{K_1}^{10}} = 13 - \frac{13}{1{,}037^{10}} = 3{,}96.$$

[1] Од ове формуле имамо користи само кад је $c > \sigma_{2k}$; за $c < \sigma_{2k}$ добићемо $y < 0$.

Дакле, овде помоћу стандардне девијације добивамо двапута већу прецизност него помоћу σ_{10}.

Ако пак узмемо $c=4{,}9$, биће $K=1{,}701$ и $K_1=1{,}24$, те је тада

$$x=8{,}51 \quad \text{и} \quad y=11{,}49,$$

тј. овде добивамо већу прецизност помоћу σ_{10} него помоћу стандардне девијације σ_2.

II.

Као што смо навели, у унутрашњости интервала $(a-c,\ a+c)$ $(c\neq 0)$ лежи најмање

$$y=n-\frac{n}{K_1^{2k}} \tag{3}$$

чланова датога низа, при чему је

$$K_1=c:\sigma_{2k}=c\cdot\left(\frac{1}{n}\sum_{\nu=1}^{n}\alpha_\nu{}^{2k}\right)^{-\frac{1}{2k}}.$$

Ако ово K_1 унесемо у (3), добићемо

$$y=n-f(k) \tag{4}$$

где је

$$f(k)=\sum_{\nu=1}^{n}\left(\frac{\alpha_\nu}{c}\right)^{2k}.$$

Ставићемо $(\alpha_\nu : c)^2=\beta_\nu$. Према претпоставци нису сви α_ν нуле; узећемо само њих у обзир. Тада, изабравши целисходну нумерацију, можемо писати

$$f(k)=\sum_{\nu=1}^{m}\beta_\nu^k, \qquad (\beta_\nu>0). \tag{5}$$

Према (4), изван затвореног интервала $(a-c,\ a+c)$ налази се мање од $f(k)$ чланова датог скупа $a_1, a_2, \ldots, a_n$.

Приметимо да је

$$f(k)>0, \qquad f(0)=m$$

и

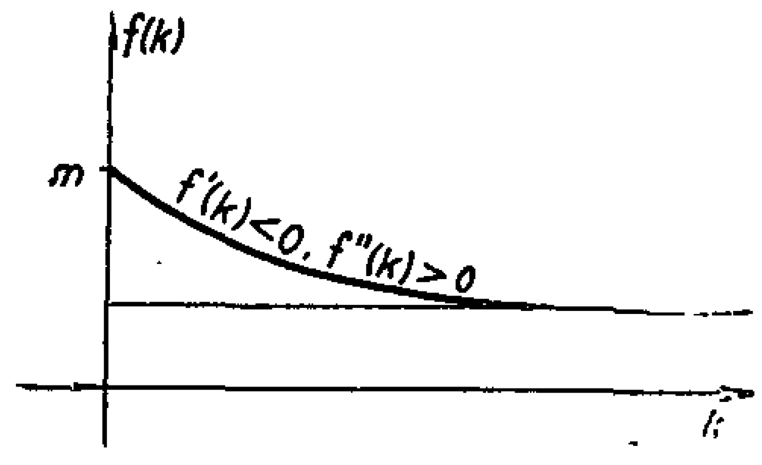
$$f'(k)=\sum_{\nu=1}^{m}\beta_\nu^k\log\beta_\nu, \qquad f'(0)=\log(\beta_1\beta_2\ldots\beta_m).$$

Како је

$$f''(k) = \sum_{\nu=1}^{m} \beta_\nu^k \log^2 \beta_\nu \qquad \text{тј.} \qquad f''(k) > 0,$$

то $f'(k)$ стално расте.

А. Ако је $f'(k)$ стално негативно, тада $f(k)$ стално опада, и како је $f(k) > 0$, то мора постојати $\lim\limits_{k \to \infty} f(k) \geqslant 0$. Отуда долазимо до закључка да ће овде прецизност бити утолико већа уколико је k веће (сл. 1).

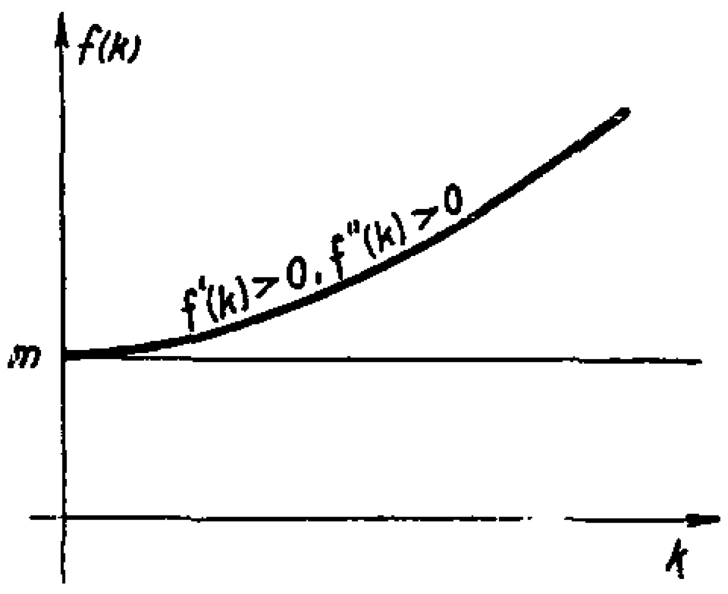


Сл. 1

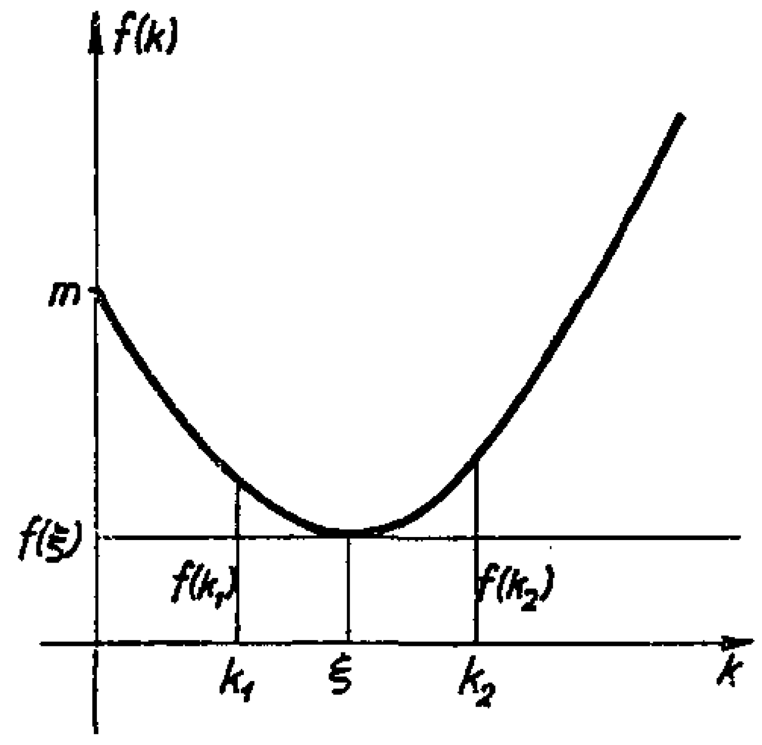
Б. Ако је $f'(k)$ стално позитивно, тада $f(k)$ стално расте, и прецизност ће утолико бити већа уколико је k мање. Имаћемо овде највећу прецизност ако је $k=1$, тј. ако узмемо стандардну девијацију (сл. 2).

В. Ако је $f'(k)$ за извесне вредности од k позитивно а за извесне негативно, тада мора постајати једно ξ за које је $f'(\xi) = 0$, тј.

$$\sum_{\nu=1}^{m} \beta_k^\xi \log \beta_\nu = 0.$$



Сл. 2



Сл. 3

Како је $f''(\xi) > 0$, то је $f(\xi) = \min$. Ако ставимо

$$k_1 = [\xi] \quad \text{и} \quad k_1 = [\xi] + 1$$

и означимо

$$f(k_0) = \min_{r=1,2} \{f(k_r)\},$$

највећу прецизност имаћемо сада за $k = k_0$ (сл. 3).

PRÉCISION DE LA DÉVIATION STANDARDE POUR UNE REPARTITION QUELCONQUE

Par B. Ivanović

L'auteur montre qu'il existe de cas où l'on peut à l'aide de la déviation standarde, arriver à un résultat plus précis qu'à l'aide de moments d'ordres supérieurs et donne l'analyse de ces cas.