F. ZÍTEK (Praga)

O PEWNYCH ESTYMATORACH ODCHYLENIA STANDARDOWEGO

Celem tej pracy jest przegląd kilku metod estymacji punktowej odchylenia standardowego w populacji generalnej o rozkładzie normalnym wraz z porównaniem ich wydajności. Praca ogranicza się do rozpatrywania tylko małych próbek oraz do jak najprostszych charakterystyk próbki. W tekście używa się elementarnych pojęć teorii estymacji, których definicje i znaczenia znajdzie czytelnik w każdym kursie statystyki matematycznej, np. w książce Craméra [1].

Zagadnienie podstawowe jest następujące: Dana jest próbka losowa $x_1, x_2, \ldots, x_n$, wzięta z populacji generalnej scharakteryzowanej gęstością prawdopodobieństwa

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\delta^2},$$

zależną od dwóch nieznanych parametrów μ, σ. Jaka funkcja owych n wartości $x_1, x_2, \ldots, x_n$ daje nam najwygodniejszą ocenę parametru σ?

Zastosowana do tego zagadnienia metoda największej wiarogodności (*maximum likelihood* [1]) daje jako estymator na σ, odchylenie standardowe próbki s określone wzorem

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2},$$

gdzie

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

oznacza średnią arytmetyczną próbki. Często bierze się zamiast s charakterystykę

$$s_1 = \sqrt{\frac{n}{n-1}}\,s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2},$$

rzekomo dlatego, żeby usunąć obciążenie estymatora s, jednak myśl ta jest błędna, gdyż w rzeczywistości oba estymatory s i s_1 są obciążone. Obciążenie to jest dosyć istotne, zwłaszcza dla próbek o małej liczności; tak na przykład dla $n=2$ jest

$$\mathrm{E}(s) = \frac{1}{\sqrt{\pi}}\,\sigma \quad \text{oraz} \quad \mathrm{E}(s_1) = \sqrt{\frac{2}{\pi}}\,\sigma.$$

Żeby naprawdę usunąć obciążenie, wystarczy pomnożyć s przez stały współczynnik będący po prostu odwrotnością jego wartości średniej podzielonej przez σ, a mianowicie

$$\sqrt{\frac{n}{2}}\cdot\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}.$$

W ten sposób otrzymamy estymator

$$\sigma_s^* = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}.$$

W tablicy 1 podano wartości współczynnika $\Gamma\left(\frac{n-1}{2}\right)\Big/\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)$.

Estymator σ_s^* jest już niemal doskonały, przynajmniej z punktu widzenia teoretycznego, jest on mianowicie nieobciążony, zgodny i asymptotycznie efektywny [1]; znamy także dokładnie jego rozkład teoretyczny. Jednak trzeba przyznać, że nasz estymator ma też pewną wadę bardzo poważną, zwłaszcza jeśli chodzi o jego zastosowanie praktyczne. σ_s^* jest bardzo skomplikowaną funkcją wartości próbki. Dwa działania nieelementarne, których wymaga jego obliczenie (w szczególności obliczanie pierwiastka kwadratowego), są już dostatecznie skomplikowane, by utrudnić, jeśli w ogóle nie uniemożliwić, stosowanie σ_s^* w praktyce. Chodzi tu zwłaszcza o statystyczną kontrolę jakości, gdy nie można żądać, by kontroler tracił więcej czasu na rachowanie, aniżeli na samą kontrolę, tym bardziej że, jak właśnie pokażemy, owe kłopoty nie są konieczne.

Tablica 1

n	$\dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)}$	$\sqrt{\dfrac{\pi}{2n(n-1)}}$	$\dfrac{\sqrt{\pi}}{n(n-1)}$	$\dfrac{1}{\sum \mathrm{E}(w_k)}$	$\dfrac{1}{\mathrm{E}(w_1)}$
2	1,25331	0,886	0,88623	0,89	0,8862
3	,79788	,512	,29541	,59	,5908
4	,62666	,362	,14770	,38	,4857
5	,53192	,280	,08862	,30	,4299
6	,46999	,229	,05908	,24	,3946
7	,42554	,193	,04220	,20	,3698
8	,39166	,167	,03165	,17	,3512
9	,36475	,148	,02462	,15	,3367
10	,34270	,132	,01969	,135	,3249
11	,32422	,120	,01611	,123	,3152
12	,30843	,109	,01343	,111	,3069
13	,29873	,100	,01136	,103	,2998
14	,28273	,093	,00974	,095	,2935
15	,27207	,086	,00844	,088	,2880

n	$\dfrac{1}{\mathrm{E}(w_2)}$	$\dfrac{1}{\mathrm{E}(w_3)}$	$\dfrac{1}{\mathrm{E}(w_4)}$	$\dfrac{1}{\mathrm{E}(w_5)}$	$\dfrac{1}{\mathrm{E}(w_6)}$	$\dfrac{1}{\mathrm{E}(w_7)}$
4	1,6					
5	1,0					
6	0,78	2,5				
7	0,66	1,4				
8	0,59	1,1	3,3			
9	0,54	0,88	1,9			
10	0,50	0,76	1,3	4,2		
11	0,47	0,68	1,1	2,3		
12	0,45	0,63	0,93	1,6	5,0	
13	0,43	0,59	0,83	1,3	2,6	
14	0,41	0,56	0,76	1,1	1,9	5,5
15	0,40	0,53	0,70	0,96	1,5	2,9

W celu usunięcia trudności rachunkowych związanych z obliczaniem pierwiastka z sumy kwadratów H. Steinhaus [7] skonstruował dyspersjometr — przyrząd do obliczania odchylenia standardowego. Ta metoda mechaniczna obliczania charakterystyki s jest zasadniczo bardzo prosta i również dokładność uzyskiwanych wyników jest praktycznie zupełnie wystarczająca. Jednak wydaje się celowe zbadać też inne metody estymacji parametru σ, korzystające z prostszych charakterystyk próbki.

Najbardziej znane spośród takich prostych charakterystyk jest *odchylenie przeciętne*

$$m = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|.$$

Wartość oczekiwana m jest

$$\mathrm{E}(m) = \sqrt{\frac{n-1}{n}}\sqrt{\frac{2}{\pi}}\sigma,$$

a więc estymatorem nieobciążonym opartym na m będzie

$$\sigma_m^* = \sqrt{\frac{\pi}{2}}\cdot\frac{1}{\sqrt{n(n-1)}}\sum_{i=1}^{n}|x_i - \bar{x}|.$$

Wartości współczynnika przy $\sum_{i=1}^{n}|x_i - \bar{x}|$ podano w tablicy 1. Widać od razu, że estymator σ_m^* jest o wiele łatwiejszy do obliczania niż σ_s^*.

Inną charakterystyką próbki, na podstawie której można estymować σ, jest *średnia różnica Gini'ego* [4]

$$g = \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}|x_i - x_j|.$$

Jej wartość oczekiwana jest

$$\mathrm{E}(g) = \frac{2}{\sqrt{\pi}}\sigma,$$

a więc

$$\sigma_g^* = \frac{\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |x_i - x_j|$$

będzie znów nieobciążonym estymatorem na σ.

Wartość współczynnika $\sqrt{\pi}/n(n-1)$ podano również w tablicy 1. W dalszym ciągu pokażemy też pewną prostą metodę obliczania sumy

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |x_i - x_j|.$$

Dotychczas rozważaliśmy trzy klasyczne miary rozsiewu wartości $x_1, x_2, \ldots, x_n$, mianowicie s, m i g. Były one znane już w zeszłym wieku, zwłaszcza astronomom i geodetom [3] (tzw. wzory Petersa i Jordana). Z rozwojem nowszych badań statystycznych wyłoniło się jeszcze kilka innych charakterystyk, tzw. *statystyk pozycyjnych*, które korzystają z uporządkowania wartości $x_1, x_2, \ldots, x_n$ według wielkości.

Oznaczmy przez $x_1^*, x_2^*, \ldots, x_n^*$ te same co poprzednio wartości z próbki losowej z populacji normalnej, uporządkowane w ciąg nierosnący

$$x_1^* \geqslant x_2^* \geqslant \ldots \geqslant x_n^*.$$

Na podstawie tego porządku określimy teraz dalsze charakterystyki próbki mogące służyć do estymacji odchylenia standardowego σ.

Najprostszą spośród nich będzie *rozstęp próbki*

$$R_1 = x_1^* - x_n^*.$$

Wyrazimy najpierw R_1 w postaci

$$R_1 = \sigma w_1.$$

Wartość oczekiwana rozstępu R będzie wówczas

$$\mathrm{E}(R_1) = \sigma\,\mathrm{E}(w_1).$$

Wartość $\mathrm{E}(w_1)$ jest dobrze znana [6]; jest to po prostu wartość oczekiwana rozstępu próbki wziętej z populacji normalnej unormowanej. Biorąc, jak zwykle, jej odwrotność (patrz tabl. 1) za współczynnik do R_1, otrzymujemy nowy nieobciążony estymator

$$\sigma^*_{R_1} = \frac{R_1}{\mathrm{E}(w_1)} = \frac{x^*_1 - x^*_n}{\mathrm{E}(w_1)},$$

bardzo prosty i łatwy do obliczania. Jeżeli, zamiast różnicy największej wartości x^*_1 i najmniejszej x^*_n, wziąć różnicę drugich z kolei wartości, to otrzymamy *drugi rozstęp*

$$R_2 = x^*_2 - x^*_{n-1}.$$

Ogólnie można napisać

$$R_k = x^*_k - x^*_{n-k+1}, \qquad k = 1, 2, \ldots, \left[\frac{n}{2}\right].$$

Wartości oczekiwane odpowiednich wielkości w_k można obliczyć na podstawie tablic Fishera-Yatesa [2]. Otrzymaliśmy więc w ten sposób ciąg estymatorów nieobciążonych

$$\sigma^*_{R_k} = \frac{x^*_k - x^*_{n-k+1}}{\mathrm{E}(w_k)}.$$

Współczynniki $1/\mathrm{E}(w_k)$ podano w dolnej części tablicy 1.

Zaletą estymatorów opartych na R_k jest ich prostota, wadą zaś fakt, że poza uporządkowaniem korzystają one z dwóch tylko wartości x_i, wskutek czego dadzą tylko małą część informacji zawartej w próbce. Żeby zapobiec temu zjawisku, można by zamiast pojedynczych rozstępów wziąć różne ich kombinacje. Najprostszym pomysłem byłoby tu wziąć średnią arytmetyczną albo gotowych już estymatorów

$$\bar{\sigma}^*_R = \frac{1}{[n/2]} \sum_{k=1}^{[n/2]} \sigma^*_{R_k},$$

albo samych rozstępów; w ten właśnie sposób otrzymamy zaproponowany przez H. Steinhausa estymator

$$\sigma^*_{St} = \frac{\sum_{k=1}^{[n/2]} R_k}{\sum_{k=1}^{[n/2]} \mathrm{E}(w_k)}.$$

Również współczynniki tego estymatora zawarte są w tablicy 1.

Sumie ΣR_k można nadać też postać

$$x^* + \ldots + x_{[n/2]} - x^*_{n-[n/2]+1} - \ldots - x^*_n,$$

lub

$$\sum_{i=1}^{n} |x_i - \nu|,$$

gdzie ν oznacza medianę próbki. Ostatnią sumę, a raczej wyrażenie,

$$\frac{1}{n} \sum_{i=1}^{n} |x_i - \nu|$$

nazywamy *przeciętnym odstępem*. Jak wynika z ogólnych twierdzeń [1], jest on zawsze niewiększy niż odchylenie przeciętne m i daje również bardziej konsekwentną miarę rozsiewu niż m.

Przypomnimy jeszcze na tym miejscu wzór von Andrae'go [3] na obliczanie średniej różnicy g. Jak łatwo się przekonać, zachodzi tu następująca równość:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |x_i - x_j| = \sum_{k=1}^{[n/2]} (n - 2k + 1) R_k.$$

Widać z tego, że średnia różnica g jest średnią ważoną rozstępów. Stosując wzór von Andrae'go średnią różnicę oblicza się łatwo i szybko.

Wobec tego, że w całej pracy ograniczamy się tylko do próbek o małej liczności ($n \leqslant 15$), nie możemy tu, niestety, wykorzystać dla estymacji σ tzw. *kwantylów próbki*, które mają właściwy sens dopiero dla większych n, chociaż, jak pokazano w [5], dają one bardzo dobre wyniki.

Pokażemy tutaj na jednym konkretnym przykładzie, w jaki sposób wykonuje się praktycznie poszczególne estymacje:

Dana jest próbka o liczności 10 i wartościach $x_i =$ 0,59 0,78 1,50 0,28 −0,24 −1,10 −1,15 1,16 0,68 −0,66.

Próbkę porządkujemy i zapisujemy według schematu

$$\begin{matrix} x_1^* & x_2^* & \ldots & x_5^* \\ x_{10}^* & x_9^* & \ldots & x_6^* \end{matrix}$$

a więc

1,50	1,16	0,78	0,68	0,59
−1,55	−1,10	−0,66	−0,24	0,28
3,05	2,26	1,44	0,92	0,31

Przez odjęcie otrzymujemy rozstępy R_k; mnożąc je teraz przez odpowiednie współczynniki dostajemy pierwszych pięć ocen

$$\begin{aligned} \sigma^*_{R1} &= 3{,}05\cdot 0{,}3249 = 0{,}99,\\ \sigma^*_{R2} &= 2{,}26\cdot 0{,}50 \quad = 1{,}13,\\ \sigma^*_{R3} &= 1{,}44\cdot 0{,}76 \quad = 1{,}09,\\ \sigma^*_{R4} &= 0{,}92\cdot 1{,}3 \quad\ = 1{,}20,\\ \sigma^*_{R5} &= 0{,}31\cdot 4{,}2 \quad\ = 1{,}30. \end{aligned}$$

Ich średnia $\bar{\sigma}^*_R = 1{,}14$.

Suma rozstępów jest 7,98, a więc $\sigma^*_{St} = 7{,}98\cdot 0{,}135 = 1{,}08$. Metodą von Andrae'go otrzymujemy

$$\sigma^*_g = 53{,}54\cdot 0{,}01969 = 1{,}05.$$

Do obliczenia m i s potrzebna jest średnia $\bar{x}$. W naszym przykładzie jest $\bar{x} = 0{,}144$. Mamy $\sum|x_i - \bar{x}| = 8{,}252$, a zatem

$$\sigma^*_m = 8{,}252\cdot 0{,}132 = 1{,}09.$$

Obliczenie s jest najtrudniejsze. Stosujemy następujący schemat:

$$\sum_{i=1}^{10} x_i^2 = 9{,}1986, \qquad 10\bar{x}^2 = 0{,}20736,$$

$$\sum_{i=1}^{10} (x_i - \bar{x})^2 = 8{,}99124, \qquad \sqrt{\sum (x_i - \bar{x})^2} = 2{,}9985.$$

Wobec tego

$$\sigma^*_s = 2{,}9985\cdot 0{,}3427 = 1{,}028.$$

Warto jeszcze zaznaczyć, że wtedy gdy $\bar{x}$ jest jednocześnie medianą, tj. gdy

$$x^*_{n/2} > \bar{x} > x^*_{n/2+1} \quad \text{dla } n \text{ parzystych}$$

lub

$$x^*_{[n/2]+1} = \bar{x} \quad \text{dla } n \text{ nieparzystych},$$

to

$$\sum_{i=1}^{n} |x_i - \bar{x}| = \sum_{k=1}^{[n/2]} R_k.$$

Korzystając z ostatniej równości możemy często obliczać σ^*_m i σ^*_{St} jednocześnie.

Zapoznaliśmy się więc z kilku estymatorami odchylenia standardowego. Wszystkie są nieobciążone, ale oczywiście nie są wszystkie jednakowo dobre. Do porównania tych estymatorów najbardziej przydałyby się ich rozkłady teoretyczne. Niestety, nie wszystkie są znane. Skoro jednak ograniczamy się do estymacji punktowej, to wystarczającą dla naszych celów miarą dobroci estymatora jest jego wariancja. W przypadku gdy chcielibyśmy obliczać przedział ufności, musielibyśmy znać rozkłady estymatorów. Wariancje są znane dla estymatorów σ_s^*, σ_g^*, σ_m^*, $\sigma_{R_1}^*$ ([3], [6]):

$$D^2(\sigma_s^*) = \left[\frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}\right)} - 1 \right] \sigma^2,$$

$$D^2(\sigma_g^*) = \frac{1}{n(n-1)} \left[\frac{\pi(n+1)}{3} + 2(n-2)\sqrt{3} - 2(2n-3) \right] \sigma^2,$$

$$D^2(\sigma_m^*) = \frac{1}{n} \left[\frac{\pi}{2} + \arcsin \frac{1}{n-1} + \sqrt{n(n-2)} - n \right] \sigma^2,$$

$$D^2(\sigma_{R_1}^*) = \frac{D^2(w_1)}{E^2(w_1)} \sigma^2.$$

Wartości wszystkich czterech współczynników przy σ^2 podano w tablicy 2. Wariancji estymatorów $\sigma_{R_k}^*$ dla $k \geqslant 2$ nie udało nam się znaleźć w dostępnym piśmiennictwie. Ich obliczanie przez całkowanie numeryczne byłoby oczywiście możliwe, ale dosyć żmudne i nie wydaje się celowe. Co do estymatora σ_{St}^*, można przypuszczać że jego wariancja nie będzie się różniła o wiele od wariancji estymatora σ_m^*. Dla lepszego porównania możemy jeszcze obliczyć minimalną wariancję regularnego estymatora na σ, wynikającą z twierdzenia Craméra-Rao. Metodą opisaną w [1] otrzymujemy

$$D^2_{\min}(\sigma^*) = \frac{1}{2n} \sigma^2$$

(patrz tabl. 2).

Tablica 2

n	$D^2(\sigma_s^*)$	$D^2(\sigma_g^*)$	$D^2(\sigma_m^*)$	$D^2(\sigma_{R_1}^*)$	$D^2_{\min}(\sigma^*)$
2	0,5708	0,5708	0,5708	0,5708	0,2500
3	,2732	,2755	,2755	,2755	,1667
4	,1781	,1803	,1848	,1826	,1250
5	,1318	,1338	,1393	,1380	,1000
6	,1045	,1062	,1119	,1120	,0833
7	,0865	,0881	,0935	,0949	,0714
8	,0738	,0752	,0803	,0830	,0625
9	,0643	,0656	,0704	,0740	,0556
10	,05700	,0581	,0626	,0670	,0500
11	,05118	,0522	,0564	,062	,0455
12	,04644	,0474	,0514	,057	,0417
13	,04250	,0434	,0471	,053	,0385
14	,03917	,0400	,0435	,050	,0357
15	,03633	,0371	,0404	,047	,0333

Z porównania wariancyj czterech estymatorów σ_s^*, σ_g^*, σ_m^*, $\sigma_{R_1}^*$ i wariancji minimalnej widać, że:

1⁰ Żaden z podanych estymatorów nie jest estymatorem efektywnym.

2⁰ Estymator σ_s^*, chociaż teoretycznie najlepszy, daje dla bardzo małych wartości n na ogół niezbyt dobre wyniki.

3⁰ σ_g^* jest prawie tak samo dobrym estymatorem jak σ_s^*. Efektywność relatywna σ_g^* w stosunku do σ_s^* jest 1,00 dla $n=2$, 0,99 dla $3 \leqslant n \leqslant 5$ oraz 0,98 dla $6 \leqslant n \leqslant 15$.

4⁰ Następny z kolei jest estymator σ_m^*. Jego efektywność jest 0,80 dla $n=10$, efektywność relatywna w stosunku do σ_s^* jest 0,90 jeszcze dla $n=15$.

5⁰ Nawet takiego prostego estymatora jak $\sigma_{R_1}^*$ wcale nie należy uważać za nie nadający się do użytku. Efektywność $\sigma_{R_1}^*$ jest dla $n=15$ równa 0,71 i efektywność relatywna w stosunku do σ_s^* jest 0,851 dla $n=10$ i 0,77 dla $n=15$. Estymator jest stosunkowo najlepszy przy $6 \leqslant n \leqslant 10$, optimum jest dla $n=7$ i $n=8$.

Najwygodniejszym estymatorem jest więc σ_g^*. Jest on prostszy od σ_g^*, nie wymaga żadnych trudnych działań arytmetycznych i daje tak samo dobre wyniki jak σ_s^*. Stosowanie estymatora σ_s^*, który jest bardzo mało wygodny z punktu widzenia praktycznego, jest więc zupełnie zbędne. To jednak wcale nie obniża wartości teoretycznej charakterystyki s, która do innych celów jest najlepsza i konieczna.

Zaletą niektórych spośród podanych estymatorów, jak np. σ_q^* jest ten fakt, że nie korzystają ze znajomości średniej $\bar{x}$. Ta właściwość może jednak w pewnym przypadku odgrywać rolę ujemną, mianowicie wtedy gdy znamy prawdziwą wartość parametru μ. Dwa tylko estymatory, σ_s^* i σ_m^*, mogą skorzystać z tej dodatkowej informacji. Kładąc μ zamiast $\bar{x}$ we wzorach na s i m otrzymujemy dwie nowe charakterystyki

$$s_0 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}, \qquad m_0 = \frac{1}{n}\sum_{i=1}^{n}|x_i-\mu|,$$

a wraz z nimi i dwa nowe nieobciążone estymatory na σ. Pierwszy z nich

$$\sigma_{s_0}^* = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{n+1}{2}\right)}\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2}$$

jest estymatorem [1], zgodnym i asymptotycznie efektywnym, o wysokiej efektywności nawet dla małych liczności n: $e_2=0{,}915$, $e_3=0{,}936$, $e_{10}=0{,}98$, $e_{14}=0{,}98$. Jego wariancja jest

$$D^2(\sigma_{s_0}^*) = \left[\frac{\Gamma\left(\frac{n+2}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)} - 1\right]\sigma^2$$

(patrz tabl. 2).

Drugim estymatorem jest

$$\sigma_{m_0}^* = \sqrt{\frac{\pi}{2}}\cdot\frac{1}{n}\sum_{i=1}^{n}|x_i-\mu|$$

z wariancją

$$D^2(\sigma_{m_0}^*) = (\pi-2)\frac{\sigma^2}{2n},$$

a wskutek tego o stałej efektywności 0,876.

Jeżeli więc znamy μ, to opłaci się stosować skomplikowaną charakterystykę s_0 i estymować za pomocą $\sigma^*_{s_0}$, bo s_0 wtedy jako wystarczająca statystyka wyzyskuje całą informację zawartą w próbce.

Prace cytowane

[1] H. Cramér, *Mathematical methods of statistics*, Princeton 1946.

[2] R. A. Fisher – F. Yates, *Statistical tables*, London 1948.

[3] Helmert, *Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers directer Beobachtungen gleicher Genauigkeit*, Astronomische Nachrichten 88 (1876), Nr 2096, S. 113.

[4] M. G. Kendall, *Advanced theory of statistics I*, London 1947.

[5] F. Mosteller, *On some useful "inefficient" statistics*, Annals of Mathematical Statistics 17 (1946), p. 377.

[6] E. S. Pearson, *The percentage limits for the distribution of range in samples from a normal population*, Biometrika 24 (1932), p. 404.

[7] H. Steinhaus, *Dyspersjometr*, Zastosowania Matematyki, ten tom, str. 321-329.

Instytut Matematyczny Polskiej Akademii Nauk

(Praca wpłynęła dnia 25. 4. 1953 r.)

Ф. ЗИТЕК (Прага).

О НЕКОТОРЫХ ОЦЕНКАХ СТАНДАРТНОГО ОТКЛОНЕНИЯ

РЕЗЮМЕ

В работе дан обзор нескольких простых оценок стандартного отклонения в генеральной совокупности с нормальным распределением. Для сравнения даны также дисперсии четырех важнейших оценок. Из сравнения следует, что некоторые, очень удобные для вычисления, оценки почти не отличаются (относительно эффективности) от обычно употребляемой довольно сложной оценки, базированой на стандартном отклонении выборки.

F. ZÍTEK (Prague)

ON CERTAIN ESTIMATORS OF STANDARD DEVIATION

SUMMARY

The paper gives a survey of several simple estimators of standard deviation in a general population with normal distribution. For the sake of comparison, the variances of four of the more important estimators are also given. The comparison shows that some estimators, very simple as regards calculation, differ very little in point of efficiency from the usually applied, fairly complicated estimator, based on the standard deviation of the sample.
