

BOOTSTRAPOWY ESTYMATOR MEDIANY DLA PRÓBY O NIEPARZYSZTEJ LICZBIE ELEMENTÓW

Joanna Kisielińska

Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych
Szkola Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: joanna_kisielinska@sggw.pl

Streszczenie: W artykule przedstawiono rozkład bootstrapowego estymatora mediany dla próby o nieparzystej liczbie elementów. Wykorzystując jego własności, oszacowano punktowo i przedziałowo medianę dla prób pochodzących z populacji o wybranych rozkładach o różnej liczebności. Przeprowadzone badania symulacyjne pokazały, kiedy lepszym oszacowaniem mediany jest mediana estymatora, a kiedy jego wartość oczekiwana. Zbadano wpływ liczebności próby na dokładność szacunków punktowych i przedziałowych.

Słowa kluczowe: bootstrapowy estymator mediany, punktowe i przedziałowe szacowanie mediany

WPROWADZENIE

W wielu zagadnieniach, nie tylko ekonomicznych, do charakterystyki poziomu badanego zjawiska nie wystarcza średnia arytmetyczna. Tak jest w przypadku, gdy rozkład cechuje asymetria. Wówczas przydatna jest inna miara średniej, czyli mediana. Jeśli badacz dysponuje danymi z całej populacji, może interesujący go parametr np. medianę obliczyć. W przypadku gdy dostępne są jedynie informacje z części populacji, może jedynie parametr oszacować.

Jedną z metod szacowania parametrów jest metoda bootstrapowa, którą stosuje się, jeśli nie jest znany rozkład cechy w badanej populacji. Warto dodać, że w zastosowaniach praktycznych zwykle rozkłady badanych cech nie są znane i jedyną informacją jaką dysponujemy jest pobrana próba losowa.

Zaproponowana przez Efrona [Efron 1979] metoda bootstrapowa polega na wtórnym próbkowaniu próby oryginalnej (pierwotnej). Jeśli metodę stosuje się w celu szacowania parametrów, dla każdej próby wtórnej wyznaczany jest intere-

sujący badacza parametr. Na podstawie zbioru obliczonych dla prób wtórnych wartości można określić bootstrapowy rozkład estymatora. Rozkład ten można wykorzystać następnie do punktowego lub przedziałowego szacowania parametru.

Podkreślić należy, że liczba wszystkich możliwych prób wtórnych jest skończona i równa n^n , gdzie n – jest liczbą elementów próby pierwotnej. Metodę bazującą na wszystkich próbach wtórnych nazywa się metodą dokładnego bootstrapu¹ (exact bootstrap method). Metodę dokładnego bootstrapu można zastosować w przypadkach gdy n nie jest zbyt duże² lub, jeśli znany jest bootstrapowy rozkład estymatora. W przypadku mediany dla próby o nieparzystej liczbie elementów „dokładny” rozkład jej bootstrapowego estymatora jest znany.

Dla próby o nieparzystej liczbie elementów estymator mediany jest estymatorem medianowo-nieobciążonym [Zieliński 2010]. Jeśli więc mediana estymatora jest różna od jego wartości oczekiwanej należy, jako oszacowanie mediany, przyjąć medianę estymatora, a nie jego wartość oczekiwaną. Z drugiej strony wiadomo, że granicznym rozkładem mediany z próby o nieparzystej liczbie elementów jest rozkład normalny o wartości oczekiwanej równej medianie rozkładu [Domański i Pruska 2000]. Z uwagi na symetryczność rozkładu granicznego, dla prób dużych zarówno mediana jak i wartość oczekiwana estymatora powinny dawać oszacowanie równie dobre.

Celem przedstawionych w niniejszym artykule badań będzie punktowe i przedziałowe szacowanie mediany bazujące na rozkładzie estymatora bootstrapowego.

Szacowanie punktowe ma na celu porównanie mediany i wartości oczekiwanej bootstrapowego estymatora mediany. Rozważania teoretyczne wskazują, że mediana powinna być oszacowaniem lepszym, natomiast wraz ze wzrostem liczebności próby różnice pomiędzy nimi powinny być coraz mniejsze.

Dla przedziałowego szacowania mediany granice przedziału ufności, ze względu na własności rozkładu estymatora bootstrapowego, mogą być z góry wyznaczone dla wszystkich prób o zadanej liczebności i zadanego poziomu ufności. Zbadany zostanie wpływ wielkości próby na dokładność oszacowań przedziałowych.

Próby wykorzystane do symulacji losowano przy użyciu generatora liczb pseudolosowych wbudowanego w program MS Excel. Wszystkie obliczenia wykonano z wykorzystaniem języka VBA.

¹ Metoda jest przedstawiona między innymi w pracy Kisielińska J. (2014) Szacowanie mediany przy użyciu dokładnej metody bootstrapowej, *Metody Ilościowe w Badaniach Ekonomicznych*. Tom XV, Nr. 3, str. 111-121.

² Dla większych liczebności próby pierwotnej czas potrzebny do wygenerowania wszystkich prób wtórnych i obliczenia dla nich interesujących parametrów może być bowiem bardzo długi.

BOOTSTRAPOWY ESTYMATOR MEDIANY DLA PRÓBY O NIEPARZYSZTEJ LICZBIE OBSERWACJI

Rozważania prowadzone będą dla próby prostej X_1, X_2, \dots, X_n wylosowanej z populacji o nieznanym rozkładzie F z medianą $M_{0.5}$. Wartości próby oznaczone zostaną jako x_1, x_2, \dots, x_n , zaś uporządkowane niemalejąco jako $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. Statystyka pozycyjna o randze k oznaczona zostanie jako $X_{(k)}$. Medianą Me z n -elementowej próby o nieparzystej liczbie elementów jest jej środkowy element:

$$Me = X_{(m)} \quad (1)$$

gdzie $m = \frac{n+1}{2}$ jest pozycją mediany.

W metodzie bootstrapowej z próby pierwotnej losowane są ze zwracaniem próby wtórne: $X_1^*, X_2^*, \dots, X_n^*$. Jeśli w próbie pierwotnej nie było powtórzeń, każda ze zmiennych X_i^* przyjmować może wartości ze zbioru $\{x_1, x_2, \dots, x_n\}$ z jednakowym prawdopodobieństwem równym $1/n$.

W próbach wtórnych medianą może być dowolny element próby pierwotnej. Jeśli bootstrapowy estymator mediany oznaczymy jako $\hat{X}_{(m)}^*$ określenie jego rozkładu wymaga obliczenia prawdopodobieństw $P(\hat{X}_{(m)}^* = x_{(l)})$, dla $l = 1, 2, \dots, n$. Prawdopodobieństwa te można wyznaczyć w sposób następujący [Maritz i Jarrett 1978, Efron 1979, Kisielińska 2014, Pekasiewicz 2015]:

$$\begin{cases} \text{dla } l = 2, \dots, n-1 \\ P(\hat{X}_{(m)}^* = x_{(l)}) = \sum_{j=0}^{m-1} \binom{n}{j} \left(\frac{l-1}{n}\right)^j \left(1 - \frac{l-1}{n}\right)^{n-j} - \sum_{j=0}^{m-1} \binom{n}{j} \left(\frac{l}{n}\right)^j \left(1 - \frac{l}{n}\right)^{n-j} \\ \text{dla } l = 1, n \\ P(\hat{X}_{(m)}^* = x_{(l)}) = \sum_{j=m}^n \binom{n}{j} \left(\frac{1}{n}\right)^j \left(\frac{n-1}{n}\right)^{n-j} \end{cases} \quad (2)$$

Wzór ten dla $l=2, \dots, n-1$ można zapisać w prostszej postaci [Maritz i Jarrett 1978]:

$$P(\hat{X}_{(m)}^* = x_{(l)}) = \frac{(n)!}{(m!)^2} \int_{(j-1)/n}^{j/n} y^m (1-y)^m dy. \quad (3)$$

Ważną cechą przedstawionego rozkładu jest to, że prawdopodobieństwa określone wzorem (2) dla poszczególnych wartości $x_{(l)}$ nie zależą od wartości obserwacji wchodzących w skład próby. Mówiąc inaczej, prawdopodobieństwa te są jednakowe dla wszystkich n elementowych prób. Łatwo pokazać ponadto, że:

$$P(\hat{X}_{(m)}^* = x_{(l)}) = P(\hat{X}_{(m)}^* = x_{(n-l+1)}). \quad (4)$$

Wartość oczekiwaną rozkładu bootstrapowego estymatora mediany dla próby o nieparzystej liczbie obserwacji można obliczyć ze wzoru:

$$E(\hat{X}_{(m)}^*) = \sum_{l=1}^n x_{(l)} \cdot P(\hat{X}_{(m)}^* = x_{(l)}), \quad (5)$$

zaś wariancję:

$$VAR(\hat{X}_{(m)}^*) = \sum_{l=1}^n (x_{(l)} - E(\hat{X}_{(m)}^*))^2 \cdot P(\hat{X}_{(m)}^* = x_{(l)}). \quad (6)$$

Wartość oczekiwana $E(\hat{X}_m^*)$ będzie równa medianie z próby pierwotnej jedynie³ w przypadku próby idealnie symetrycznej względem środkowego elementu, czyli takiej, która spełnia warunek:

$$x_{(m)} - x_{(l)} = x_{(n-l+1)} - x_{(m)}. \quad (7)$$

Na wylosowanie reprezentatywnej próby symetrycznej są szanse jedynie wówczas, gdy rozkład, dla którego pobrano próbę jest rozkładem symetrycznym. Pamiętać jednak należy, że mediana jest parametrem interesującym jedynie dla rozkładów asymetrycznych.

Znając rozkład bootstrapowego estymatora mediany można zbudować przedziały ufności. Ponieważ rozkład ten jest dyskretny, jako granice przedziału warto przyjąć elementy próby. Dla zadanego poziomu ufności $1-\alpha$, lewą granicą przedziału ufności $Me_{\frac{\alpha}{2}}$ dla wszystkich n -elementowych prób o nieparzystej liczbie

elementów będzie wówczas największy element $x_{(d)}$, który spełnia warunek:

$$F(x_{(d)}) \leq \frac{\alpha}{2}, \quad (8)$$

gdzie F jest dystrybucją estymatora \hat{X}_m^* . Granicą prawą przedziału $Me_{1-\frac{\alpha}{2}}$ natomiast będzie najmniejszy element $x_{(g)}$, który spełnia warunek:

$$F(x_{(g)}) \geq 1 - \frac{\alpha}{2}. \quad (9)$$

Podkreślić należy, że jeśli próba nie spełnia warunku (7) przedziały te będą asymetryczne względem mediany z próby, a wobec tego jako miarę dokładności oszacowania lepiej używać szerokości przedziału, a nie jego połowy.

WYNIKI OBLICZEŃ

Eksperymenty symulacyjne przeprowadzono dla populacji o sześciu rozkładach. Rozkłady te oznaczone zostały jako:

- $F(10,7)$: rozkład F o liczbie stopni swobody licznika 10 i mianownika 7,

³ Nie można oczywiście wykluczyć wylosowania próby niesymetrycznej, dla której mediana będzie równa wartości oczekiwanej, będzie to jednak przypadek, a nie reguła.

- $F(5,2)$: rozkład F o liczbie stopni swobody licznika 5 i mianownika 2,
- $CHI(3)$: rozkład χ^2 o liczbie stopni swobody 3,
- $CHI(10)$: rozkład χ^2 o liczbie stopni swobody 10,
- $N(5,1)$: rozkład normalny,
- $N(5,4)$: rozkład normalny.

Z każdej populacji wylosowano po 100 prób o liczebnościach $n = 21, 31, 41, 51, 61, 71, 81, 91, 101, 111, 121$. W celu większej porównywalności wyników dla różnych rozkładów losowano liczby losowe z przedziału $[0,1]$, które traktowano następnie jako dystrybuantę rozkładu. Dla każdego n obliczono bootstrapowy rozkład mediany, a następnie dla wszystkich prób wyznaczono parametry rozkładu: medianę i wartość oczekiwaną oraz przedziały ufności.

Tabela 1. Średnie odległości oszacowań mediany od faktycznej mediany rozkładu wyznaczone ze 100 prób

n	$F(10,7), M_{0,5}=1,030$		$F(5,2), M_{0,5}=1,252$		$CHI(3), M_{0,5}=2,366$	
	$ d_1 $	$ d_2 $	$ d_1 $	$ d_2 $	$ d_1 $	$ d_2 $
21	0,059711	0,078324	0,192693	0,304717	0,155731	0,193178
31	0,044853	0,028693	0,078128	0,149936	0,054525	0,082834
41	0,021680	0,031027	0,078977	0,128069	0,054326	0,074151
51	0,015578	0,021089	0,054239	0,086433	0,039551	0,050352
61	0,009630	0,021438	0,037836	0,082665	0,023370	0,052524
71	0,001856	0,009608	0,018773	0,051427	0,001571	0,019940
81	0,001284	0,009143	0,014879	0,046507	0,000635	0,019574
91	0,001247	0,006294	0,016100	0,037439	0,000183	0,012125
101	0,013141	0,014929	0,008399	0,019089	0,034271	0,036979
111	0,002381	0,007478	0,013762	0,034142	0,004663	0,016976
121	0,007537	0,009489	0,026806	0,039185	0,018990	0,022568
n	$CHI(10), M_{0,5}=9,342$		$N(5,1), M_{0,5}=5$		$N(5,4), M_{0,5}=5$	
	$ d_1 $	$ d_2 $	$ d_1 $	$ d_2 $	$ d_1 $	$ d_2 $
21	0,059711	0,078324	0,192693	0,304717	0,155731	0,193178
31	0,044853	0,028693	0,078128	0,149936	0,054525	0,082834
41	0,021680	0,031027	0,078977	0,128069	0,054326	0,074151
51	0,015578	0,021089	0,054239	0,086433	0,039551	0,050352
61	0,009630	0,021438	0,037836	0,082665	0,023370	0,052524
71	0,001856	0,009608	0,018773	0,051427	0,001571	0,019940
81	0,001284	0,009143	0,014879	0,046507	0,000635	0,019574
91	0,001247	0,006294	0,016100	0,037439	0,000183	0,012125
101	0,013141	0,014929	0,008399	0,019089	0,034271	0,036979
111	0,002381	0,007478	0,013762	0,034142	0,004663	0,016976
121	0,007537	0,009489	0,026806	0,039185	0,018990	0,022568

Uwagi: d_1 - odległość mediany estymatora bootstrapowego od mediany rozkładu,
 d_2 - odległość wartości oczekiwanej estymatora bootstrapowego od mediany rozkładu.

Źródło: badania własne

W Tabeli 1 przedstawione zostały średnie odległości oszacowań mediany od jej faktycznej wartości, wyznaczone dla sześciu populacji ze 100 prób. Odległość oznaczoną jako d_1 obliczono dla oszacowania w postaci mediany z prób bootstrapowych, zaś d_2 dla oszacowania w postaci wartości oczekiwanej. Dla rozkładów asymetrycznych ($F(10,7)$, $F(5,2)$, $CHI(3)$, $CHI(10)$) mediana była oszacowaniem lepszym od wartości oczekiwanej (jedynie w czterech przypadkach na 44 było przeciwnie). Dla dwóch rozkładów symetrycznych ($N(5,1)$, $N(5,4)$) lepszym oszacowaniem była zwykle wartość oczekiwana (z wyjątkiem trzech przypadków na 22).

Ponieważ dla rozkładów asymetrycznych mediana estymatora jest lepszym oszacowaniem niż wartość oczekiwana, stosowanie wariancji jako miary dokładności nie będzie odpowiednie. Można użyć w tym celu odległości międzykwartylowej [Zieliński 2010] lub szerokości przedziałów ufności.

W przypadku prób z populacji o rozkładach F , systematyczne zmniejszanie odległości zaobserwować można dla rozkładu $F(10,7)$ do $n = 91$, zaś dla rozkładu $F(5,2)$ do $n = 101$. Dalsze zwiększanie próby w niektórych przypadkach daje oszacowanie gorsze niż dla próby mniejszej. Najmniejsze różnice między oszacowaniem mediany, a jej faktyczną wartością uzyskano dla próby 91 elementowej w przypadku rozkładu $F(10,7)$ i 101 elementowej dla rozkładu $F(5,2)$. Na błędy oszacowań wpływ ma wariacja rozkładu populacji. Rozkład $F(5,2)$ ma większą wariancję niż rozkład $F(10,7)$, co skutkuje większymi błędami (zarówno dla mediany jak i wartości oczekiwanej bootstrapowego rozkładu estymatora) niemal dla wszystkich liczebności prób.

Dla populacji o rozkładzie χ^2 najlepsze oszacowanie mediany uzyskano dla $n = 91$ w przypadku rozkładu $CHI(3)$ oraz dla $n = 111$ w przypadku rozkładu $CHI(10)$. Zwiększanie liczebności próby w początkowym zakresie n skutkowało zmniejszeniem błędów. Dla wyższych wartości n oszacowania były w niektórych przypadkach gorsze niż dla n mniejszych. Rozkład $CHI(10)$ ma większą wariancję niż $CHI(3)$ i oszacowanie były dla niego zwykle gorsze (z wyjątkiem $n = 101$ i 111).

Najlepsze oszacowanie mediany dla obydwu populacji o rozkładzie normalnym uzyskano dla $n = 111$. Błędy oszacowań dla $n=121$ były w niektórych przypadkach większe nawet kilkukrotnie. Podobnie jak w przypadku populacji o rozkładach asymetrycznych, błędy oszacowań były większe dla populacji o większej wariancji.

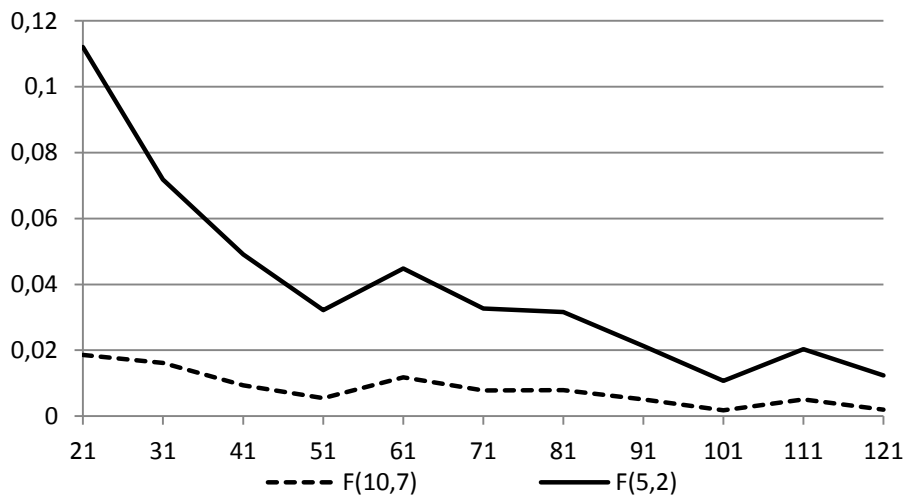
Wzrost błędów oszacowań mediany rozkładu wraz ze wzrostem liczebności próby (dla większych n) stoi w sprzeczności z teorią i wynika z błędów numerycznych występujących podczas obliczania prawdopodobieństw zgodnie ze wzorem (2) dla większych prób. Błędy te występują mimo zastosowania w programie komputerowym podwójnej precyzji dla zmiennych rzeczywistych. Oznacza to, że zwiększanie liczebności próby, przy aktualnej dokładności komputerów niekoniecznie musi prowadzić do zwiększania dokładności szacunku.

Wpływ wielkości próby na różnicę pomiędzy medianą i wartością oczekiwaną estymatora bootstrapowego dla prób wylosowanych z populacji o rozkładach F , χ^2 i normalnym przedstawiono na Rysunkach 1, 2 i 3. Odległości te obliczono jako różnice pomiędzy wartościami średnimi wyznaczonymi ze 100 prób.

W przypadku prób z populacji o rozkładach F , można zaobserwować niemal systematyczne zmniejszanie różnic między medianą i wartością oczekiwaną. Zmniejszanie różnic występuje także w przypadku rozkładu χ^2 , choć nie jest już systematyczne. Dla prób z populacji o rozkładach normalnych wielkość próby nie wpływa na analizowane różnice.

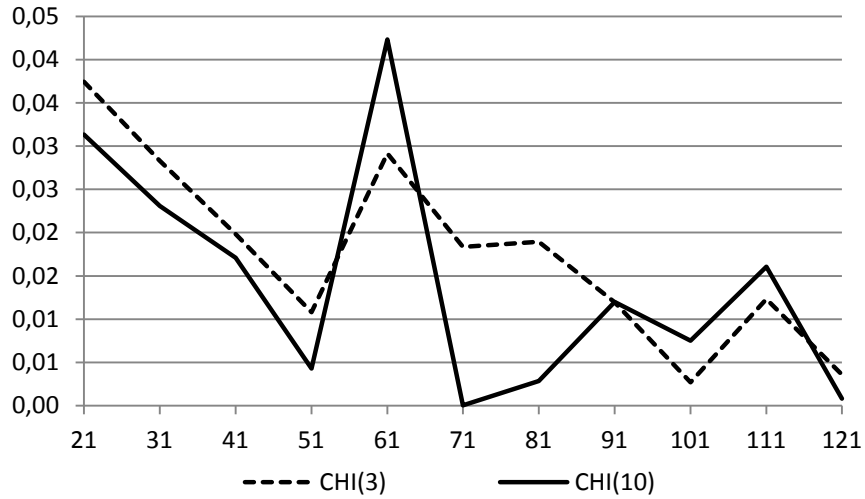
Różnicę pomiędzy medianą i wartością oczekiwaną estymatora bootstrapowego są większe dla prób wylosowanych z populacji o większej wariancji w przypadku rozkładów F i normalnych. Dla rozkładu χ^2 takiej prawidłowości nie zaobserwowano.

Rysunek 1. Wartości bezwzględne różnic pomiędzy medianą i wartością oczekiwaną bootstrapowego estymatora mediany dla prób z populacji o rozkładach F



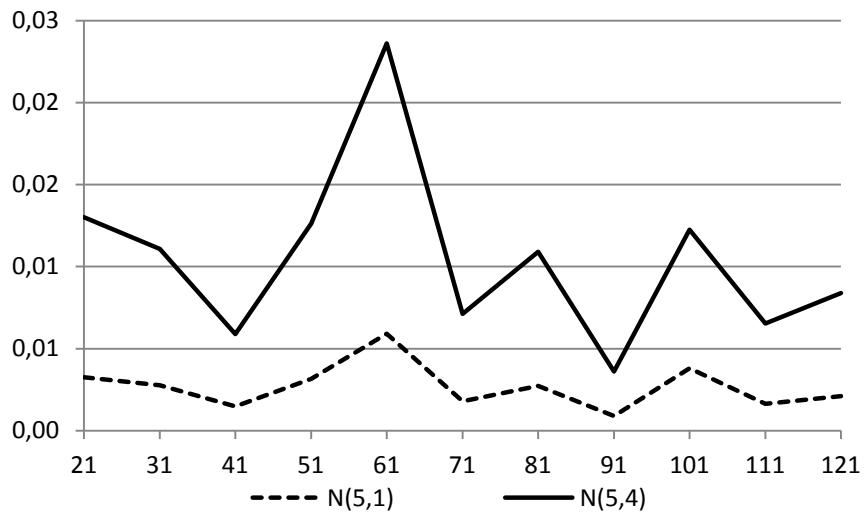
Źródło: badania własne

Rysunek 2. Wartości bezwzględne różnic pomiędzy medianą i wartością oczekiwaną bootstrapowego estymatora mediany dla prób z populacji o rozkładach χ^2



Źródło: badania własne

Rysunek 3. Wartości bezwzględne różnic pomiędzy medianą i wartością oczekiwaną bootstrapowego estymatora mediany dla prób z populacji o rozkładach normalnych



Źródło: badania własne

Znając rozkład bootstrapowego estymatora mediany można wyznaczyć przedziały ufności. Granice przedziałów spełniać muszą warunki (8) i (9). Ponieważ prawdopodobieństwa poszczególnych realizacji estymatora są jednakowe dla wszystkich prób o zadanej liczebności, granice te wyznaczają elementy próby o określonej pozycji w uporządkowanej próbie. W Tabeli 2 podano pozycje granic dla poziomu ufności $1-\alpha = 0,95$ dla prób liczących $n = 21, 31, 41, 51, 61, 71, 81, 91, 101, 111$ i 121 . Podkreślić należy, że w przypadku stosowania metody percentyli przedstawionej między innymi w [Pekasiewicz 2015] losowanie prób wtórnych nie jest potrzebne do szacowania mediany i innych kwantyli.

Na Rysunku 4 przedstawiono wpływ wielkości próby na średnią szerokość przedziałów ufności dla $1-\alpha = 0,95$ dla badanych sześciu rozkładów wyznaczoną ze 100 prób.

Szerokość przedziałów ufności zmniejsza się niemal równomiernie wraz ze wzrostem liczebności próby. Pozwala to wnioskować, że dla mniejszych i większych elementów próby (odpowiadających dystrybucie na poziomie 0,05 i 0,95) rozkład obliczony jest z mniejszymi błędami numerycznymi niż dla obserwacji środkowych.

Wariancja rozkładu w populacji wpływa na szerokości przedziałów ufności. Im większa wariancja, tym przedziały są oczywiście szersze. Najszersze przedziały obserwujemy dla przypadku rozkładu $CHI(10)$, a następnie kolejno dla rozkładów $N(5,4)$, $CHI(3)$, $F(5,2)$, $N(5,1)$ i $F(10,7)$.

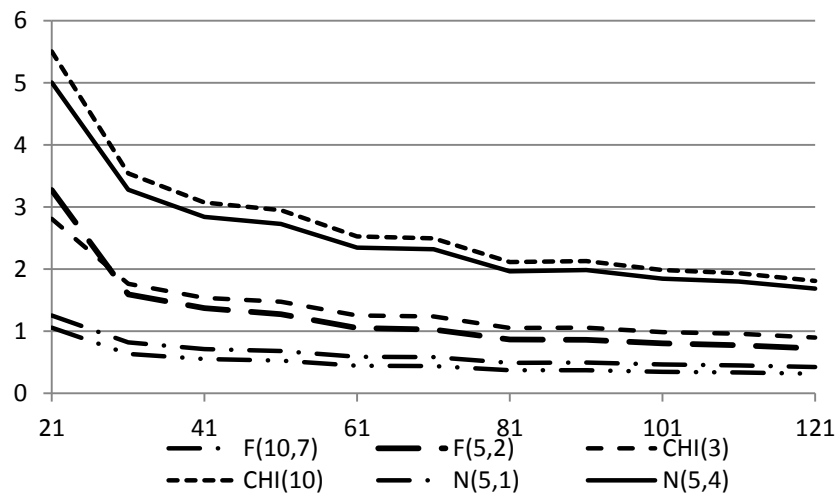
Tabela 2. Numery obserwacji będące granicami przedziałów ufności dla $1-\alpha = 0,95$

n	21	31	41	51	61	71	81	91	101	111	121
$n_{(d)}$	6	11	15	19	24	28	33	37	42	46	51
$n_{(g)}$	16	21	21	33	38	44	49	55	60	66	71

Uwagi: $n_{(d)}$ – pozycja lewej granicy przedziału ufności,
 $n_{(g)}$ – pozycja prawej granicy przedziału ufności.

Źródło: badania własne

Rysunek 4. Wpływ wielkości próby na średnią szerokość przedziałów ufności dla $1-\alpha = 0,95$



Źródło: badania własne

PODSUMOWANIA

Przeprowadzone badania pozwalają na wyciągnięcie następujących wniosków:

- W przypadku punktowego szacowania mediany, dla prób z populacji o wziętych pod uwagę rozkładach asymetrycznych, mediana bootstrapowego estymatora mediany jest oszacowaniem lepszym niż wartość oczekiwana. Jest to zgodne z teorią, ponieważ mediana jest estymatorem medianowo-nieobciążonym. Dla rozkładów symetrycznych (normalnych) lepszym oszacowaniem okazała się wartość oczekiwana. Wraz ze wzrostem liczebności próby różnice pomiędzy medianą i wartością oczekiwaną estymatora bootstrapowego malały dla prób ze wszystkich populacji.
- Zwiększanie liczebności próby w zakresie początkowych wartości n prowadziło do lepszych oszacowań mediany. Dla większych n natomiast wzrost liczebności próby w wielu przypadkach prowadził do zwiększania błędów. Jest to prawdopodobnie skutkiem błędów numerycznych podczas obliczania prawdopodobieństw dla poszczególnych realizacji estymatora dla większych prób. Przy aktualnej dokładności komputerów wzrost liczebności próby niekoniecznie musi prowadzić do zwiększania dokładności szacunku.

- Ponieważ prawdopodobieństwa poszczególnych realizacji bootstrapowego estymatora mediany zależą jedynie od liczebności próby, pozycje granic przedziałów ufności są jednakowe dla wszystkich prób o zadanej liczebności. Z tego też względu losowanie prób wtórnych do ich wyznaczenia nie jest potrzebne.
- Dla zadanego poziomu ufności szerokości przedziałów ufności zmniejszają się wraz ze wzrostem liczebności próby.

BIBLIOGRAFIA

- Efron B. (1979) Bootstrap methods: another look at the jackknife, *The Annals of Statistics*, Vol. 7, No. 1, 1-26.
- Domański C., Pruska K. (2000) *Nieklasyczne metody statystyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Kisielińska J. (2014) Szacowanie mediany przy użyciu dokładnej metody bootstrapowej, *Metody Ilościowe w Badaniach Ekonomicznych*. Tom XV, Nr. 3, str. 111-121.
- Maritz J. S., Jarrett R. G. (1978) A note on estimating the variance of the sample median, *Journal of the American Statistical Association*, Vol. 73, No. 361, 194-196.
- Pekasiewicz D. (2015) *Statystyki pozycyjne w procedurach estymacji i ich zastosowania w badaniach ekonomicznych*, Wydawnictwo Uniwersytetu Łódzkiego.
- Zieliński R. (2010) O średniej arytmetycznej i medianie. *Matematyka Stosowana*, Tom XI/52. http://www.matstos.pjwstk.edu.pl/no11/no11_zielinski.pdf

MEDIAN BOOTSTRAP ESTIMATOR FOR AN ODD SAMPLE

Abstract: The article presents the distribution of the median bootstrap estimator for the sample with an odd number of elements. Using its properties, made point and interval estimation of median for samples taken from a population of selected distributions of different sizes. The conducted simulation studies have shown, when the median of estimator is a better estimate, and when its expected value. It was examined the impact of sample size on the accuracy of point estimates and intervals.

Keywords: bootstrap median estimator, point and interval estimation of median