Grażyna Trzpiot
Uniwersytet Ekonomiczny w Katowicach

WYBRANE STATYSTYKI ODPORNE

Wprowadzenie

Obserwacje oddalone (*outliers*) są takimi obserwacjami w próbie, które mogą powodować zakłócenia w ocenie relacji w próbie. Nie jest to termin o znaczeniu pejoratywnym; obserwacje oddalone mogą być poprawne, ale powinny być identyfikowane dla oceny błędów. Poczynając od 60., zaproponowano wiele metod silnych i odpornych *(robust* and *resistant)* mniej wrażliwych na obserwacje oddalone. Mogą one konkurować, a nawet wygrywać ze standardowymi statystycznymi metodami. Omawiana tematyka jest przedmiotem wcześniejszych prac autorki zawsze w kontekście zastosowań w ekonomii (Trzpiot 2009, 2011a, 2011b). Artykuł ten ma charakter opisowy w powiązaniu z przygotowywanym podręcznikiem.

1. Statystyki jednowymiarowe położenia i skali

Średnia z próby może być załamana przez pojedynczą obserwację. Jeżeli dowolna obserwacja ma wartość taką, że $y_i \to \pm\infty$, wówczas średnia z próby $\bar{y} \to \pm\infty$, w przeciwieństwie do mediany z próby, która nie jest wrażliwa na pojedyncze wartości zmierzające do nieskończoności. Mówimy, że mediana jest odporna na duże błędy, podczas gdy średnia nie. Faktycznie mediana może znieść do 50% dużych błędów zanim będzie arbitralnie duża; mówimy, że ma *punkt załamania* 50-proc., podczas gdy dla średniej mamy odpowiednio 0%.

Średnia jest efektywnym estymatorem parametru położenia dla rozkładu normalnego, dlatego może być wykorzystywana jako estymator parametru położenia dla rozkładów zbliżonych do normalnych. Metody odporne powinny mieć wysoką efektywność w otoczeniu zakładanego modelu statystycznego.

Dlaczego nie jest wystarczające przesianie danych i odrzucenie obserwacji odstających? Należy rozważyć wiele aspektów metodologicznych:
1. Praktycy, nawet eksperci statystycy, nie zawsze przeglądają zbiory danych.
2. Ostre decyzje, czy zachować, czy odrzucić obserwacje mogą być niezbyt trafne. Proponujemy nadać wagi wątpliwym obserwacjom. Możemy również odrzucić kompletnie złe obserwacje.

3. Może być zadaniem trudnym lub wręcz niemożliwym umiejscowienie obserwacji odstających w wielowymiarowym lub mocno zrestrukturyzowanym zbiorze danych.
4. Odrzucenie obserwacji odstających wpływa na rozkład teoretyczny (zmiennej losowej), który musi być skorygowany. W szczególności wariancja będzie niedoszacowania w „wyczyszczonym” zbiorze.

Dla ustalonego rozkładu definiujemy *relatywną efektywność* estymatora. Efektywność estymatora $\hat{\theta}$ względem innego estymatora $\tilde{\theta}$ możemy zmierzyć, posługując się następującą miarą efektywności:

$$RE(\tilde{\theta},\hat{\theta}) = \frac{D^2(\hat{\theta})}{D^2(\tilde{\theta})} \tag{1.1}$$

Granicę $RE(\tilde{\theta},\hat{\theta})$ przy rosnącej do nieskończoności wielkości próby nazywamy *efektywnością asymptotyczną*:

$$ARE(\tilde{\theta},\hat{\theta}) = \lim_{n\to\infty} RE(\tilde{\theta},\hat{\theta}) \tag{1.2}$$

Estymatorem asymptotycznie najefektywniejszym jest estymator, którego asymptotyczna efektywność równa się jedności. Można problem zdefiniować również w odniesieniu do asymptotycznych wariancji. Jeżeli estymator $\hat{\theta}$ nie jest znany, wówczas zakładamy, że jest efektywnym estymatorem. Pojawiają się trudności z obciążonymi estymatorami, których wariancja jest mała lub wynosi zero.

Proponowanym w literaturze rozwiązaniem jest wykorzystanie błędów średniokwadratowych, innym – przeskalowanie $\theta/E(\hat{\theta})$. Iglewicz (1983) proponuje wykorzystanie wariancji logarytmu estymatora $\hat{\theta}$: $D^2(\log\hat{\theta})$ jako estymatora parametru skali*. Zastosujmy podejście *ARE* do oceny średniej i mediany (Venables, Ripley, 2002). Dla rozkładu normalnego

$$ARE(\text{mediana, średnia}) = \frac{D^2(\textit{średnia})}{D^2(\textit{mediana})} = 2/\pi \approx 64\%$$

Dla rozkładów o innych wartościach rozkładów w ogonach mediana ma lepsze własności. Przykładowo, dla rozkładu t-Studenta z pięcioma stopniami swobody, a to jest często rozkład zgodny z rozkładem błędów modeli, *ARE* (mediana, średnia) $\approx$ 96%.

* jest niezależna od skali

Kolejny przykład podał Tukey (1960). Zakładamy, że mamy n obserwacji $Y_i \sim N(\mu, \sigma^2)$ dla $i = 1,..., n$ oraz chcemy estymować wartość wariancji σ^2. Rozważmy dwa estymatory $\hat{\sigma}^2 = s^2$ oraz $\tilde{\sigma}^2 = d^2\pi/2$, gdzie:

$$d = \frac{1}{n}\sum_i \left|Y_i - \overline{Y}\right|$$

oraz stała jest wybrana tak, że dla rozkładu normalnego $d \to \sqrt{2/\pi}\sigma$. Wówczas $ARE(\tilde{\sigma}^2, s^2) = 0,876$.

Załóżmy, że dla każdego Y_i mamy obserwacje z rozkładu $N(\mu, \sigma^2)$ z prawdopodobieństwem $1 - \varepsilon$ oraz wartości z rozkładu $N(\mu, 9\sigma^2)$ z prawdopodobieństwem ε. Zauważmy, że obydwie wariancje dla wszystkich obserwacji oraz wariancja niezakłóconego rozkładu obserwacji są proporcjonalne do σ^2. Otrzymujemy dane zawarte w tab. 1.

Tabela 1

Wartości *ARE* dla wybranych wartości ε

ε (%)	$ARE(\tilde{\sigma}^2, s^2)$
0	0,876
0,1	0,948
0,2	1,016
1	1,44
5	2,04

Źródło: Na podstawie (Venables, Ripley, 2002).

Mieszanka rozkładów z zakłóceniem $\varepsilon = 1\%$ jest nieodróżnialna od rozkładu normalnego, zwłaszcza w praktycznych zastosowaniach, dlatego optymalność s^2 jest bardzo wrażliwa. Mówimy o braku *odporności efektywności* estymatora.

Znajdujemy odmienne estymatory parametru σ niż $d\sqrt{\pi/2}$ (mają punkt załamania 0%). Dwa proponowane rozwiązania przyjmowane jako estymatory są porównywalne:

$$IQR = X_{(3n/4)} - X_{(n/4)} \tag{1.3}$$

oraz

$$MAD = \underset{i}{\text{mediana}} \{|Y_i - \underset{j}{\text{mediana}}(Y_j)|\} \tag{1.4}$$

Przykładowo, dla rozkładu normalnego otrzymujemy odpowiednio następujące wyniki:

$MAD \rightarrow$ mediana $\{|Y - \mu|\} \approx 0{,}6745\sigma$,
$IQR \rightarrow \sigma[\Phi-1(0{,}75) - \Phi-1(0{,}25)] \approx 1{,}35\sigma$

Obydwa estymatory są efektywne, ale bardzo odporne na obserwacje oddalone w zbiorze danych. Dla estymatora MAD i dla rozkładu normalnego mamy ARE= 37% (Staudte, Sheather, 1990, s. 123).

W kolejnym kroku rozważań zakładamy, że mamy n niezależnych obserwacji Y_i z rodziny z parametrem położenia o funkcji gęstości $f(y-\mu)$, oraz funkcja f jest symetryczna względem zera. Zatem μ jest wartością centralną (mediana, średnia, jeżeli istnieje) dystrybuanty Y_i. Rozważamy rozkład niewiele różniący się od rozkładu normalnego. Mamy wiele estymatorów wartości μ. Wśród tego zbioru estymatorów znajdujemy średnią z próby, medianę z próby i estymatory wyznaczane metodą największej wiarygodności (MNW). Dodatkowo rozpatrujemy *średnią uciętą*, która jest średnią dla $1-2\alpha$ wartości rozkładu, zatem αn obserwacji jest usuniętych z każdego końca badanego rozkładu (największych i najmniejszych).

2. M-estymatory parametrów położenia i skali

Rozważymy jako estymatory parametru położenia znane z literatury *M--estymatory*. Nazwa pochodzi od sformułowania „prawie" MNW estymatory (‘*MLElike' estimators*).

Analizując funkcję gęstości f, możemy zdefiniować funkcję $\rho = -\log f$. Wówczas estymator największej wiarygodności wyznaczamy jako:

$$\min_{\mu} \sum_{i} -\log f(y_i - \mu) = \min_{\mu} \sum_{i} \rho(y_i - \mu) \tag{2.1}$$

Powyższe przekształcenie jest użyteczne, jeżeli funkcja ρ nie jest funkcją gęstości. Zapiszmy, jako $\psi = \rho'$ (jeżeli ta pochodna istnieje), wówczas otrzymujemy:

$$\sum_{i} \psi(y_i - \hat{\mu}) = 0 \text{ lub } \sum_{i} w_i(y_i - \hat{\mu}) = 0 \tag{2.2}$$

gdzie:
$w_i = \psi(y_i - \hat{\mu})/(y_i - \hat{\mu})$.

To sugeruje iteracyjne metody rozwiązania, przy czym wagi uaktualniamy przy każdej kolejnej iteracji.

Przykłady M-estymatorów

Średnia z próby odpowiada funkcji $\rho(x) = x^2$, mediana z próby odpowiada funkcji $\rho(x) = |x|$. Dla dowolnego n mediana jest rozwiązaniem zapisanego problemu.

Funkcja

$$\psi(x) = \begin{cases} x, & |x| < c \\ 0, & |x| \geq c \end{cases}$$

odpowiada *uciętej metryce*; duże odległości pomiędzy wartościami nie mają żadnego wpływu.

Funkcja*

$$\psi(x) = \begin{cases} -c, & x < -c \\ x, & |x| < c \\ c, & x > c \end{cases}$$

odpowiada metryce Winsora i obejmuje wartości ekstremalne obserwacji jako $\mu \pm c$.

Odpowiednia funkcja $\rho = -\log f$ jest następująca:

$$\rho(x) = \begin{cases} x^2, & |x| < c \\ c(2|x| - c), & |x| \geq c \end{cases}$$

i wyznacza funkcję gęstości o rozkładzie Gaussa w centrum rozkładu, mającą podwójnie wykładnicze ogony. Ten estymator zdefiniował Huber (1981). Zauważmy, że jeżeli $c \to 0$, w granicy otrzymujemy medianę, oraz jeżeli $c \to \infty$, wówczas granicą jest średnia. Wartość $c = 1{,}345$ zapewnia 95% efektywności dla rozkładu normalnego.

Funkcja podwójnie ważąca Tukey'a ma postać:

$$\psi(t) = t[1 - \left(\frac{t}{R}\right)^2]_+^2$$

gdzie $[.]_+$ oznacza dodatnie wartości. To jest, jak zwykło się określać, „łagodne" (*soft*) ucinanie. Wartość $R = 4{,}685$ zapewnia 95% zgodności efektywności dla rozkładu normalnego (Venables, Ripley, 2002).

Kolejny przykład to funkcja ψ Hampela (1986), która jest kawałkami liniowa:

---

* Pojęcie określone przez Charlesa P. Winsora (por. Dixon, 1960).

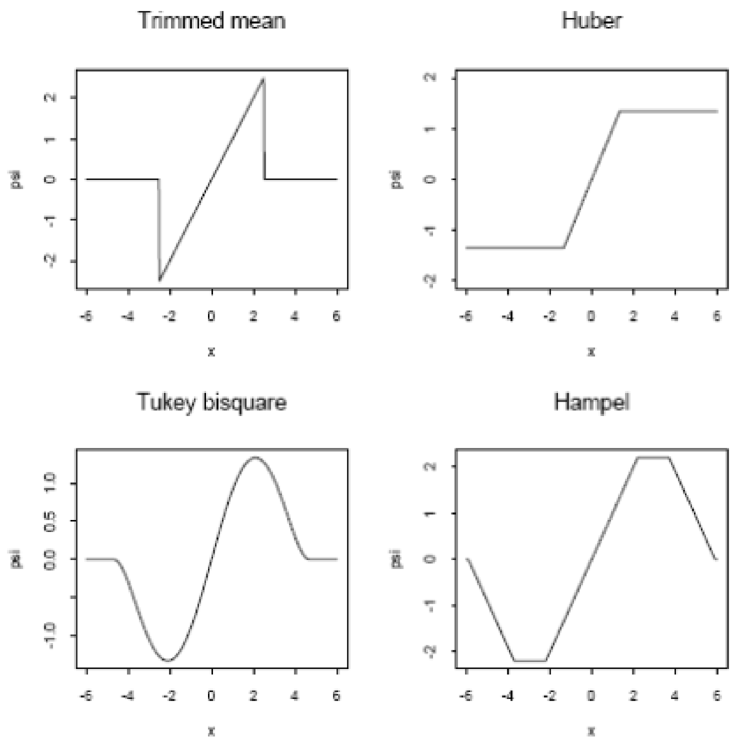
$$\psi(x)=\begin{cases}|x|, & 0<|x|<a\\ a, & a<|x|<b\\ a(c-|x|)/(c-b), & b<|x|<c\\ 0, & c<|x|\end{cases}$$

Ilustracja omówionych estymatorów (rys. 1) wymagała przyjęcia umownych wartości parametrów: $a = 2{,}2s$, $b = 3{,}7s$, $c = 5{,}9s$. Zauważamy oczywiście problem skali. W czterech ostatnich przypadkach mamy współczynniki skali (c, R lub s).

Możemy zastosować estymator do przeskalowania rezultatów:

$$\min_{\mu} \sum_{i} \rho\left(\frac{y_i-\mu}{s}\right) \tag{2.3}$$

dla współczynnika skali s, przykładowo estymator MAD. Alternatywnie, możemy estymować s w podobny sposób.



Rys. 1. Przykłady funkcji ważących dla M-estymatorów

Wykorzystując MNW dla gęstości $s^{-1}f((x - \mu)/s)$, otrzymujemy równanie

$$\sum_i \psi\left(\frac{y_i - \mu}{s}\right)\left(\frac{y_i - \mu}{s}\right) = n \tag{2.4}$$

które nie jest odporne (oraz obciążone dla rozkładu normalnego). Trzeba to równanie zmodyfikować do

$$\sum_i \chi\left(\frac{y_i - \mu}{s}\right) = (n-1)\gamma \tag{2.5}$$

dla ograniczonej funkcji χ, gdzie γ jest wybierane tak, aby uzyskać zgodność z rozkładem normalnym, zatem $\gamma = E\chi(N)$.

Przykładem niech będzie następna propozycja Hubera:

$$\chi(x) = \psi(x)^2 = \min(|x|, c)^2 \tag{2.6}$$

W bardzo małych próbach należy skupić uwagę dodatkowo na zmienności $\hat{\mu}$ w przypadku zastosowania metryki Winsora (Huber, 1981). Jeżeli położenie μ jest znane, możemy zastosować ten estymator, zastępując $n - 1$ przez n, celem estymacji jedynie współczynnika skali s.

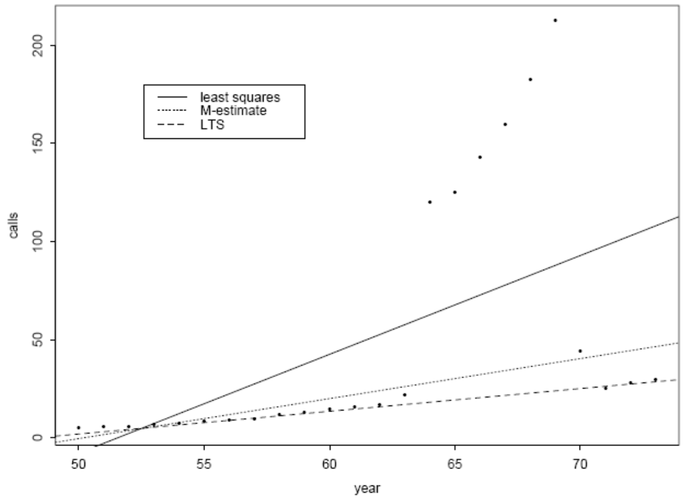
3. Własności modeli regresji

Omówimy koncepcję *odpornej regresji* w zakresie modeli liniowych, która mówi o niezagrażających zachowaniach w bieżących niewłaściwych wartościach danych. W terminologii, którą wprowadzimy, *odporna regresja* ma wysoki punkt załamania – proponujemy 50%.

Rozważymy zamianę metody najmniejszych kwadratów (MNK) przez jeden dwóch z metod:

1. LMS – najmniejsze medianowe kwadraty: minimalizują medianę kwadratów reszt. Bardziej ogólnie LQS – minimalizuje pewien kwantyl (przykładowo 80%) kwadratów reszt.
2. LTS – najmniejsze ucięte kwadraty: minimalizują sumę kwadratów najmniejszych q reszt. Oryginalnie q zawiera trochę powyżej 50%.

Omówione podejścia wymagają znacznie więcej obliczeń numerycznych niż metoda najmniejszych kwadratów, ponieważ nie mamy zachowanej różniczkowalności. Obydwie metody wychwytują efekt wielowymiarowych obserwacji oddalonych i koncentrują się na dobrym dopasowaniu, do co najmniej powyżej 50% danych. W konsekwencji są mniej efektywne w przypadku braku obserwacji odstających (LMS bardziej niż LTS). Aby zilustrować pewne problemy, rozważmy przykład. Rousseeuw i Leroy (1987) rozpatrują roczne dane liczby połączeń telefonicznych w Belgii (rys. 2). Zaprezentowano liniową funkcję regresji (MNK – least squares), regresję z M-estymacją oraz najmniejsze ucięte kwadraty reszt (LTS).



Rys. 2. Miliony połączeń telefonicznych w Belgii, 1950-1973

Źródło: (Rousseeuw, Leroy, 1987).

Linia LQS jest następująca: $\hat{y}$ = 56,16 +1,16 t (rok). Wykonane badania pokazują, że dla lat 1964–1969 powinna być badana całkowita długość połączeń (w minutach) w miejsce liczby połączeń (jak to było wykonywane w latach 1963-1970).

4. Odporna regresja*

Regresję odporną zdefiniowano w latach 80. XX w. (Huber, 1981). Pierwsza najbardziej znana regresja była określona następująco:

$$\min_{b} \underset{i}{mediana} |y_i - x_i b|^2$$

jako najmniejsze medianowe kwadraty (*least median of squares* – LMS).

Uzasadnieniem dla kwadratów reszt jest następująca obserwacja, gdy n jest parzyste, wówczas wybierana jest mediana. To jest bardzo odporna metoda regresji, dodatkowo nie wymaga estymacji parametru skali. Jest jednak bardzo nieefektywna pokrywa, co najwyżej $1/\sqrt[3]{n}$ danych. Dodatkowo, cechuje się wrażliwością na obserwacje centralne w zbiorze danych (Hettmansperger, Sheather, 1992; Davies, 1993, *§*2.3).

---

* *Resistant regression*.

Rousseeuw (1987) sugeruje regresję najmniejszych uciętych kwadratów (*least trimmed squares* – LTS):

$$\min_{b} \sum_{i} \left| y_i - x_i b \right|^2_{(i)} \tag{4.1}$$

Ta metoda jest bardziej efektywna, ale oddziela same krańcowe obserwacje. Rekomendowana suma kwadratów reszt nie powinna przekraczać wartości $q = [(n + p + 1)/2]$.

Następnie wprowadzono *S-estymatory*, dla których współczynniki równania są wyznaczane jako rozwiązanie zadania

$$\sum_{i=1}^{n} \chi\left(\frac{y_i - x_i b}{c_o s}\right) = (n-p)\beta \tag{4.2}$$

z najmniejszym parametrem skali s. W równaniu (4.2) jako funkcja χ jest zazwyczaj przyjmowana całkowalna podwójnie ważąca funkcja Tukey'a

$$\chi(u) = \begin{cases} u^6 - 3u^4 + 3u^2, |u| \le 1 \\ 1, \qquad\qquad\quad |u| \ge 1 \end{cases}$$

Wartości $c_0 = 1{,}548$ i $\beta = 0{,}5$ są wyznaczane celem spełnienia warunku zgodności, jeżeli rozkład błędów jest rozkładem normalnym. To daje efektywność 28,7% przy rozkładzie normalnym, która jest niska, ale lepsza niż LMS i LTS.

Jedynie w kilku specjalnych przypadkach (LMS dla jednowymiarowej regresji ze stałą) możemy ten problem optymalizacyjny rozwiązać dokładnie, wykorzystując aproksymacyjne metody kolejnych przybliżeń (Marazzi, 1993). Wiele tych metod wykorzystuje podejście metody najmniejszych kwadratów, proponując dopasowanie najmniejszych kwadratów dla wybranych q punktów ze zbioru danych. Następne losowo sprawdzają duże próby dla tego dopasowania.

5. Mocna regresja

W modelu regresji mamy dwa podstawowe źródła błędów: wartości obserwacji y_i oraz odpowiadający wektor p* wartości zmiennych objaśniających $\boldsymbol{x}_i$ (*regressors*). Większość metod regresji rozważa jedynie pierwszy rodzaj źródła błędów. W pewnych okolicznościach (przykładowo planowanie eksperymentów) błędy zmiennych objaśniających mogą być ignorowane. Tak jest w przypadku M-estymatorów, którymi zajmiemy się w tym punkcie.

* n obserwacji $(y_i, x_{i1}, \ldots, x_i p)$

Rozważmy problem regresji dla n przypadków (y_i, $\boldsymbol{x}_i$) z modelu

$$y = \boldsymbol{x}\beta + \varepsilon \tag{5.1}$$

dla p-wymiarowego wektora $\boldsymbol{x}$.

M-estymatory

Przyjmujemy skalowanie dla funkcji gęstości $f(e/s)/s$ dla ε oraz przyjmujemy $\rho = -\log f$, wówczas estymator maksymalnej wiarygodności minimalizuje

$$\sum_{i=1}^{n} \rho\left(\frac{y_i - x_i b}{s}\right) + n \log s \tag{5.2}$$

Załóżmy, że s jest znane oraz funkcja $\psi = \rho'$. Wówczas w MNW, wyznaczając b celem estymacji β, rozwiązujemy nieliniowe równanie:

$$\sum_{i=1}^{n} x_i \psi\left(\frac{y_i - x_i b}{s}\right) = 0 \tag{5.3}$$

Zapiszmy reszty jako: $r_i = y_i - \boldsymbol{x}_i b$. Rozwiązanie równania (5.3) lub minimalizacja względem (5.2) definiuje M-estymatory względem współczynników β.

Znaną metodą rozwiązania (5.3) jest metoda iteracyjna ważonych najmniejszych kwadratów, z wagami określonymi następująco:

$$w_i = \psi\left(\frac{y_i - \mu}{s}\right) / \left(\frac{y_i - \mu}{s}\right) \tag{5.4}$$

Iteracja jest zbieżna jedynie dla wypukłych (*convex*) funkcji ρ oraz dla niemalejących (Tukey, 1960), a równanie (5.3) może mieć wiele pierwiastków. W takich przypadkach należy wybrać dobry punkt startowy i uważnie przeprowadzić iterację.

W zastosowaniach współczynnik skali s jest nieznany. Łatwy i odporny estymator współczynnika skali to MAD (względem pewnego przyjętego centrum). Można go zastosować dla reszt bliskich zero, również dla pozostających w pewnym w otoczeniu albo dla reszt z odpornego dopasowania. Alternatywnie, możemy estymować s, wykorzystując „prawie" MNW estymatory (MLE-*like way*). Znajdując punkt stacjonarny równania (5.2) względem s, otrzymujemy:

$$\sum_{i} \psi\left(\frac{y_i - x_i b}{s}\right)\left(\frac{y_i - x_i b}{s}\right) = n \tag{5.5}$$

Rozwiązanie nie jest odporne oraz obciążone dla rozkładu normalnego (Venables, Ripley, 2002).

W przypadku jednowymiarowym możemy to równanie zmodyfikować przekształcając do postaci:

$$\sum_i \chi\left(\frac{y_i - x_i b}{s}\right) = (n-p)\gamma \tag{5.6}$$

MM-estymacja

Możliwe jest połączenie odporności oraz efektywności M-estymatorów. Takim rozwiązaniem jest MM-estymator zaproponowany przez Yohai, Stahel i Zamar (1991)*. MM-estymator to M-estymator, który wykorzystuje współczynniki wyznaczone na pierwszym etapie przez S-estymator oraz stały współczynnik skali dany przez S-estymator. To pozwala uzyskać (dla $c > c_0$) wysoki punkt załamania S-estymatorów oraz wysoką efektywność dla rozkładu normalnego. Przy znacznych kosztach obliczeń otrzymujemy to, co najlepsze z obydwu omówionych podejść.

Podsumowanie

W przedstawionym artykule omówiono wybrane statystyki odporne podstawowych parametrów wraz z ich własnościami. W szczególności omówiono wybrane estymatory parametrów położenia i skali. Zwrócono uwagę na podstawowe uwarunkowania odpornej regresji. Stosując klasyczne estymatory, nie wracamy do założeń, które towarzyszą metodom wyznaczania tych estymatorów. Brak spełnienia tych założeń powoduje trudności w wyznaczaniu rozwiązań formułowanych zadań. Estymatory odporne wymagają zastosowania metod przybliżonych, iteracyjnych. Celem efektywnego wyznaczenia wartości tych estymatorów ważne jest spojrzenie na własności numeryczne metod iteracyjnych stosowanych do rozwiązań zapisanych zadań. Wiele programów komputerowych wspomagających procesy analizy danych, takich jak S-Plus czy *Statistica* lub SAS, mają funkcje powiązane ze statystycznymi metodami odpornymi. Badanie porównawcze efektywności tych metod iteracyjnych jest odrębnym zadaniem powiązanym ze statystyką odporną.

Bibliografia

Davies P.L. (1993): *Aspects of Robust Linear Regression*. „Annals of Statistics", 21, s. 1843-1899.

Dixon W.J. (1960): *Simplified Estimation for Censored Normal Samples*. „Annals of Mathematical Statistics", 31, s. 385-391.

* (Zob. Marazzi, 1993).

Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986): *Robust Statistics. The Approach Based on Influence Functions*. John Wiley and Sons, New York.

Hettmansperger T.P., Sheather S.J. (1992): *A Cautionary Note on the Method of Least Median Squares*. „American Statistician" 46, s. 79-83

Huber P.J. (1981): *Robust Statistics*. John Wiley and Sons, New York.

Iglewicz B. (1983): Robust *Scale Estimators and Confidence Intervals for Location*. W: *Understanding Robust and Exploratory Data Analysis*. Eds. D.C. Hoaglin, F. Mosteller, J.W. Tukey. John Wiley and Sons, New York, s. 405-431.

Marazzi A. (1993): *Algorithms, Routines and* S *Functions for Robust Statistics*. Wadsworth and Brooks/Cole. Pacific Grove, CA.

Rousseeuw, P. J., Leroy, A.M. (1987): *Robust Regression and Outlier Detection*. John Wiley and Sons, New York.

Staudte R.G., Sheather S.J. (1990): *Robust Estimation and Testing*. John Wiley and Sons, New York.

Trzpiot G. (2009): *Extreme Value Distributions and Robust Estimation*. Acta Universitatis Lodziensis. Folia Economica 228, Łódź, s. 85-92.

Trzpiot G. (2011a): *Odporna analiza szeregów czasowych*. Prace Naukowe nr 165, 171-179 Uniwersytet Ekonomiczny, Wrocław.

Trzpiot G. (2011b): *Wybrane odporne metody estymacji beta*. W: *Modelowanie preferencji a ryzyko '11*. Red. T. Trzaskalik. Wydawnictwo Uniwersytetu Ekonomicznego, Katowice, s. 133-148.

Trzpiot G. (2012): *Odporna regresja kwantylowa*. W: *Dylematy ekonometrii 2*. Red. J. Biolik. Wydawnictwo Uniwersytetu Ekonomicznego, Katowice, s. 147-158.

Tukey J.W. (1960): *A Survey of Sampling from Contaminated Distributions*. W: *Contributions to Probability and Statistics*. Eds I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, H. Mann. Wiley and Sons, New York.

Yohai V., Stahel W.A., Zamar R.H. (1991): *A Procedure for Robust Estimation*. John Wiley and Sons, New York.

Venables W.N., Ripley B.D. (2002): *Modern Applied Statistics with S-PLUS*. Springer-Verlag.

SOME ROBUST STATISTICAL METHODS

Summary

Outliers are sample values that cause surprise in relation to the majority of the sample. This is not a pejorative term; outliers may be correct, but they should always be checked for transcription errors. Many *robust* and *resistant* methods have been developed since 1960 to be less sensitive to outliers. This methods can be used instead or be even better than classical one. Robust methods were used early in me works (Trzpiot 2009, 2011a, 2011b) as an application in finance and economy. This article has a descriptive character, connected with new book for students.