



Tomasz Szkutnik

Uniwersytet Ekonomiczny w Katowicach
Wydział Zarządzania
Katedra Ekonometrii
tomasz.szkutnik@ue.katowice.pl

**WPLYW LEWOSTRONNEGO CENZUROWANIA
DOTKLIWOŚCI STRAT PRZY SZACOWANIU
RYZYKA OPERACYJNEGO NA PRZYKŁADZIE
ROZKŁADU BURRA TYPU III**

Streszczenie: W artykule poruszono problem wykorzystania różnych sposobów raportowania w odniesieniu do szacowania ryzyka operacyjnego metodą LDA. Zaawansowane metody pomiaru w najnowszych rozwiązaniach Komitetu Bazylejskiego ds. Nadzoru Finansowego prawdopodobnie nie będą już służyły do określania wielkości kapitału regulacyjnego, jednak ich przydatność dalej może stanowić jeden z elementów systemu zarządzania ryzykiem w instytucji finansowej. Celem artykułu jest zbadanie wpływu utraty informacji o zdarzeniach operacyjnych na stabilność szacunków ryzyka metodą LDA. W tym kontekście proponowane w artykule rozwiązania wykorzystujące rozkłady lewostronnie cenzurowane mogą stanowić dodatkowy kierunek ich rozwoju. W artykule przedstawiono problem na przykładzie lewostronnie cenzurowanego rozkładu Burra typu III, który może być podstawą do wyznaczenia rozkładu dotkliwości strat. Rozważana koncepcja polegająca na cenzurowaniu danych została porównana z powszechnie wykorzystywanym w praktyce rozwiązaniem polegającym na ich ucinaniu. Przeprowadzone badanie symulacyjne pokazało, że w przypadku cenzurowania danych można oczekiwać bardziej stabilnych wyników, które będą bliższe sytuacji pełnej wiedzy o zachodzących procesach, na podstawie których jest szacowane ryzyko operacyjne. Estymacja parametrów z zastosowaniem danych cenzurowanych wykorzystywana powszechnie w innych obszarach badawczych może być jednym z kierunków rozwoju również w przypadku ryzyka operacyjnego w instytucjach finansowych.

Słowa kluczowe: ryzyko operacyjne, LDA, lewostronne cenzurowanie danych, rozkład Burra III.

Wprowadzenie

Jednym z najczęściej wskazywanych problemów związanych z szacowaniem wielkości ryzyka operacyjnego w bankach jest dostępność danych. Z uwagi na specyfikę poruszanego problemu, jakim jest zarządzanie ryzykiem operacyjnym, jak również na wykorzystywane w tym zakresie metody jego pomiaru, to od przyjęcia wstępnych założeń dotyczących otoczenia i sposobu funkcjonowania instytucji finansowych zależą często ostateczne szacunki ryzyka. W początkowych pracach związanych z pomiarem wielkości narażenia na ryzyko operacyjne były przyjmowane uproszczone założenia w ramach wykorzystywanych modeli. W szczególności proces generujący dane był postrzegany w najprostszych kategoriach, bez uwzględniania specyfiki powstawania, rejestracji i wykorzystywania informacji o stratach operacyjnych banku.

Komitet Bazylejski ds. Nadzoru Finansowego (w skrócie BCBS) w chwili obecnej rewiduje swoje stanowisko dotyczące pomiaru i zarządzania w zakresie ryzyka operacyjnego. W harmonogramie prac na lata 2015-2016 BCBS skupiał się głównie wokół czterech zasadniczych obszarów, do których należy m.in. aspekt zapewnienia odpowiedniej równowagi pomiędzy prostotą, porównywalnością i wrażliwością na ryzyko w całej strukturze regulacyjnej [BIS, 2014]. W ramach prowadzonych prac BCBS kontynuuje program reform postkryzysowych, do których należy przede wszystkim przywrócenie zaufania do wskaźników kapitałowych. Jedną z tego pochodnych jest wskazana przez BCBS konieczność rewizji i zmiana standardowych podejść do szacowania ryzyka w instytucjach finansowych. W odniesieniu do ryzyka operacyjnego BCBS w dokumencie skierowanym do konsultacji pod koniec 2014 [por. BIS, 2014] pierwszy raz poddaje krytyce dotychczasowe metody pomiaru w tym obszarze oraz informuje, że metody zaawansowane (w skrócie AMA) najprawdopodobniej nie będą wymagane przy szacowaniu kapitału regulacyjnego. Wraz z pojawieniem się kolejnego dokumentu konsultacyjnego w marcu 2016, jakim jest „Standardised Measurement Approach for Operational Risk” [BIS, 2016], wiadomo, że w najbliższym czasie nastąpi wyeliminowanie metod AMA i zastąpienie ich ujednoliconym rozwiązaniem zwanym SMA (Standardised Measurement Approach).

BCBS uważa, że modelowanie ryzyka operacyjnego dla celów kapitału regulacyjnego jest nadmiernie skomplikowane. Ponadto według BCBS metody AMA spowodowały nadmierną zmienności aktywów ważonych ryzykiem oraz niedostateczny poziom kapitału regulacyjnego w odniesieniu do niektórych banków [BIS, 2016].

Brak konieczności szacowania kapitału regulacyjnego metodami AMA nie musi koniecznie oznaczać końca wypracowanych przez ostatnie lata rozwiązań.

Niektórzy wręcz sugerują, że upadek AMA oznaczający koniec wykorzystania modeli wewnętrznych do celów kapitałowych może spowodować uwolnienie możliwości analitycznych w celu opracowania bardziej przydatnych modeli wewnętrznych w zakresie zarządzania ryzykiem operacyjnym niż w przypadku AMA [PWC, 2015].

Pomijając bieżące prace BCBS, które skupiają się na metodach podstawowych i szacowaniu kapitału regulacyjnego, artykuł skupia się na metodach pomiaru należących dotychczas do tzw. metod zaawansowanych (w skrócie AMA, szerzej o sposobach pomiaru ryzyka operacyjnego zob. w Chernobai i in. [2007]), które mogą znaleźć zastosowanie w przypadku szacowania kapitału ekonomicznego, jak również stanowić nowe narzędzie benchmarku czy metodę testowania warunków skrajnych.

Celem artykułu jest wskazanie problemu poprawnego określania procesu generującego dane będącego w rzeczywistości pewną formą cenzurowania informacji. W ramach proponowanej koncepcji możliwe jest uwzględnienie specyfiki procesu cenzurowania już na etapie estymacji parametrów odpowiednich rozkładów prawdopodobieństwa. Cenzurowanie danych wpisuje się tu w ogólne procesy związane z powstawaniem, rejestracją oraz wykorzystaniem informacji o stratach banku, które mogą pochodzić zarówno z samej sprawozdawczości w instytucji finansowej, jak i z zewnętrznych źródeł informacji. Cenzurowanie informacji może być również pewną formą kompromisu pomiędzy rejestrowaniem informacji jedynie powyżej arbitralnie określonych progów a pełnym raportowaniem strat operacyjnych.

Artykuł skupia się na metodzie rozkładu strat (w skrócie LDA) jako jednej z najczęściej wykorzystywanych wewnętrznych metod pomiaru. Punktem wyjścia w badaniu jest określenie hipotetycznych procesów generujących dane i późniejsza ocena ich oddziaływania na wielkości szacowanego ryzyka metodą LDA. Przeprowadzone tu badanie symulacyjne ma na celu pokazanie, jak cenzurowanie informacji wpływa na niestabilność szacunków ryzyka przy wykorzystaniu różnych wariantów metody LDA.

1. Metoda LDA

W metodzie LDA wielkość ryzyka jest wyznaczana ze zagregowanego rozkładu strat, który powstaje z dwóch procesów cząstkowych (tj. procesu dotkliwości strat oraz procesu częstości strat). Podział taki daje możliwość uwzględnienia wielu dodatkowych informacji na etapie estymacji parametrów odpowiednich ciągłych lub dyskretnych rozkładów prawdopodobieństwa. Dzięki temu jest rów-

niez możliwa ocena wrażliwości końcowej wielkości ryzyka w odniesieniu do ilości i rodzaju dodatkowych informacji, jakie mogą być wykorzystane w metodzie LDA. W artykule pominięto formalny opis metody LDA, jest on bowiem identyczny jak w pracach Chernobai i in. [2007] oraz Klugman i in. [2008]. Wyznaczenie zagregowanego rozkładu strat opiera się na metodzie Monte Carlo, która wraz ze szczegółowym opisem jest zawarta w takich pracach, jak Szkutnik, Basiaga [2013] czy Shevchenko [2011].

Jak zostało wspomniane na wstępie, punktem wyjścia w badaniu jest rozpatrzenie różnych procesów generowania danych. W odniesieniu do ryzyka operacyjnego są możliwe trzy schematy będące szczególnymi przypadkami cenzurowania informacji. Należy w tym miejscu podkreślić fakt, że cenzurowanie danych, w zależności od jego formy, w pierwszej kolejności będzie w różny sposób oddziaływać na cząstkowe procesy z metody LDA. W praktyce wpływ ten zostanie w dalszej kolejności przeniesiony na zagregowany rozkład strat. Niepełna informacja o próbie w istotny sposób zaburza szacowaną częstość zjawiska. Ponadto brak części informacji o wielkościach strat może mieć istotny wpływ także na ocenę rozkładu dotkliwości strat. Wpływ ten jest trudny do zweryfikowania w ogólnym przypadku. Bez rozpatrzenia szczególnych wariantów związanych z wybranymi rozkładami prawdopodobieństwa lub procesami stochastycznymi dla cząstkowych procesów niemożliwe wydaje się podanie jakichkolwiek wielkości opisujących ogólny wpływ cenzurowania informacji na wielkość czy stabilność szacunków ryzyka.

Aby mieć częściową wiedzę o wpływie zbierania i przetwarzania informacji o zdarzeniach i stratach operacyjnych w instytucji finansowej na wybrany sposób modelowania, należy problem ten postrzegać przez pryzmat celu, jaki stoi za pomiarem ryzyka. Największe straty słusznie są utożsamiane z realnym zagrożeniem dla instytucji i stąd zainteresowanie wszystkich metod kwantyfikacji ryzyka skupia się na dokładnym pomiarze tego zakresu dziedziny. Pominięcie na poziomie zbierania informacji lub budowy modeli strat operacyjnych o niskiej dotkliwości, a wysokiej częstości wynika zazwyczaj ze względów praktycznych. Zbyt duże obciążenie związane z koniecznym nakładem pracy dotyczącym raportowania strat o teoretycznie niskiej przydatności dla końcowego modelu może być niekiedy nieefektywne (z punktu widzenia celu, jakim jest zarządzanie ryzykiem w skali globalnej). W tym zakresie warto wspomnieć o innym wariantcie związanym ze zbieraniem informacji, jakim jest ich cenzurowanie.

Brak możliwości zarejestrowania zdarzeń lub zbieranie częściowych informacji jest znane w praktyce w takich obszarach, jak badania środowiskowe, kliniczne, analiza czasu trwania, czy w zagadnieniach statystyki medycznej. Proces

cenzurowania danych w tych dziedzinach jest naturalny i wynika ze specyfiki badanych zjawisk. Cenzurowanie danych informuje, że zdarzenie zaszło, jednak dokładna wielkość straty nie została zarejestrowana wcale lub w części.

2. Klasyfikacja danych cenzurowanych

Z uwagi na złożoność problemu, jakim jest cenzurowanie danych, nie istnieje jeden sposób ich klasyfikacji. Warto wskazać w tym zakresie podział przedstawiony w pracy Millard [2013, s. 175], gdzie cenzurowanie to:

- ucinanie lub cenzurowanie,
- lewostronne, prawostronne, obustronne cenzurowanie,
- pojedyncze lub wielokrotne (progresywne) cenzurowanie,
- cenzurowanie typu I lub typu II.

W pierwszym podpunkcie widać pewną nieścisłość, która informuje, że cenzurowanie dzieli się na cenzurowanie lub ucinanie. Niemniej jednak jest to poprawne, gdyż ogólne określenie cenzurowania dotyczy tych dwóch bardziej szczegółowych przypadków. Pojęcie cenzurowania można tu rozumieć w ogólnym wymiarze jako proces częściowej utraty informacji. Ponadto jako jedna z jego szczególnych form jest nazywana tak samo, jednak dotyczy bardziej szczegółowego przypadku rejestracji danych.

Ucinanie obserwacji występuje wtedy, gdy część danych powyżej lub poniżej progów ucinania jest nierejestrowana i brak jest informacji o zajściu zdarzenia. Natomiast w przypadku drugim, tj. cenzurowania danych, informacje powyżej lub poniżej progu cenzurowania są rejestrowane w kategoriach wystąpienia i zliczane. W tej sytuacji pełne raportowanie dotyczy pozostałej dziedziny danych.

W artykule rozważono tylko najprostszy przypadek cenzurowania danych, w którym proces cenzurowania jest procesem lewostronnego jednokrotnego cenzurowania typu I. Oznacza to, że pełne raportowanie i pomiar danych następuje tylko powyżej określonego progu „ c ”. Poniżej tego progu strata operacyjna jest zliczana bez pełnej informacji o zdarzeniu (w szczególności bez informacji o finansowej wielkości straty czy o dokładnym momencie zajścia zdarzenia). Ponadto pojedyncze cenzurowanie oznacza, że próg „ c ” jest stały dla całego zbioru danych. Ponadto cenzurowanie typu I informuje, że liczba cenzurowanych obserwacji jest zmienną losową i nie jest znana z góry (w przypadku cenzurowania typu II liczba niecenzurowanych obserwacji jest znana, a badanie kończy się w momencie osiągnięcia określonej liczby zdarzeń).

W niniejszym artykule proces częstości występowania strat jest rozumiany jako proces liczący, którego kolejne realizacje są niezależnymi zmiennymi losowymi.

wymi o takim samym dyskretnym rozkładzie prawdopodobieństwa. Nie zakłada się tu sytuacji, w której proces częstości występowania strat będzie wykazywał odmienną dynamikę w zależności od momentu czasu, w jakim zaszło zdarzenie. W tym kontekście przyjęto również założenie, że cenzurowanie lub ucinanie danych będzie wpływało tylko i wyłącznie na samą częstość zdarzeń bez zmiany charakteru czy dynamiki procesu stochastycznego, jak również bez zmiany wykorzystanego rozkładu prawdopodobieństwa do modelowania tego obszaru.

Cenzurowanie danych może w największym stopniu zaburzać modelowanie procesu dotkliwości strat operacyjnych. Cenzurowanie zmienia całkowicie postać funkcji wiarygodności, jaka służy do oszacowania parametrów wybranych ciągłych rozkładów prawdopodobieństwa. Brak uwzględnienia w procesie estymacji faktu modyfikacji pierwotnych danych może mieć duży wpływ na postać końcowego rozkładu.

3. Ogólna postać funkcji wiarygodności w przypadku różnych typów cenzurowania lub ucinania obserwacji

Prezentowana funkcji wiarygodności zakłada, że wielkość progów, poniżej lub powyżej którego dane będą poddane procesowi ucinania lub cenzurowania, jest niezależna od wartości zmiennej losowej. Ogólna postać funkcji wiarygodności w omawianym przypadku wygląda następująco [Klein, 2003, s. 74]:

$$L = \prod_{i \in D} f(x_i) \cdot \prod_{i \in R} S(C_r) \cdot \prod_{i \in L} (1 - S(C_l)) \cdot \prod_{i \in I} [S(L_i) - S(R_i)] \quad (1)$$

Relacja dana wzorem (1) przedstawia ogólny problem, jaki może być rozpatrzony dla różnych form cenzurowania lub ucinania danych. Wielkości oznaczają odpowiednio: $f(\cdot)$ – funkcja gęstości; $F(\cdot)$ – dystrybuanta; $S(x) = P(X > x)$ – funkcja przeżycia.

4. Rozkład Burra typu III

Za rozkład dotkliwości strat został przyjęty rozkład Burra typu III. Jak podaje Feroze i in. [2015], wybór ten jest często podyktowany względami praktycznymi, ponieważ elastyczność i możliwość dopasowania, jakie dają rozkłady z rodziny Burra (łącznie 12 rozkładów), powodują, że często jest to rozsądny wybór w modelowaniu zjawisk o nieznanym a priori rozkładzie. Rozkład Burra

typu III może stanowić alternatywę dla częściej wykorzystywanych rozkładów, takich jak rozkład Weibulla czy rozkład Burra typu XII.

Funkcja gęstości rozkładu Burra typu III jest dana jako:

$$f(x) = \alpha \cdot \beta x^{-\alpha-1} (1 + x^{-\alpha})^{-\beta-1} \quad (2)$$

Odpowiednio dystrybuanta tego rozkładu to:

$$F(x) = (1 + x^{-\alpha})^{-\beta} \quad (3)$$

Bazując na ogólnym wzorze (1) dla funkcji wiarygodności przy danych lewostronnie cenzurowanych, postać funkcji wiarygodności w tym przypadku można zapisać jako:

$$L(\alpha, \beta) = (1 + x_{(r+1)}^{-\alpha})^{-r\beta} \cdot (\alpha\beta)^{n-r} \cdot \prod_{i=r+1}^n x_{(i)}^{-\alpha-1} \cdot (1 + x_{(i)}^{-\alpha})^{-\beta-1} \quad (4)$$

Odpowiednio logarytm z funkcji wiarygodności danej wzorem (4) to:

$$l(\alpha, \beta) = -r\beta \cdot \ln(1 + x_{(r+1)}^{-\alpha}) + (n - r) \ln \alpha\beta + \\ - (\alpha + 1) \sum_{i=r+1}^n \ln(x_{(i)}) - (\beta + 1) \cdot \\ \cdot \sum_{i=r+1}^n \ln(1 + x_{(i)}^{-\alpha}) \quad (5)$$

gdzie przez wielkość k jest rozumiana liczba obserwacji cenzurowanych, a oznaczenie n to liczba wszystkich obserwacji.

5. Konstrukcja badania symulacyjnego

Ocena wpływu procesów ucinania oraz cenzurowania na wielkości ryzyka może być analizowana w dwóch wymiarach. W pierwszej kolejności analiza dotyczy wpływu utraty informacji na proces dotkliwości strat jako jeden z fundamentów metody LDA. W drugim etapie został zbadany wpływ cenzurowania danych na końcową wielkość ryzyka z modelu rozkładu strat.

W badaniu przyjęto, że proces częstości występowania strat to jednorodny proces Poissona (z ustalonym arbitralnie parametrem λ). Głównym celem badania symulacyjnego jest pokazanie skali potencjalnych różnic w ocenie wielkości ryzyka dla jednego tylko procesu modelowanego metodą LDA.

Za proces dotkliwości strat przyjęto wspomniany w poprzednich częściach pracy rozkład Burra typu III, ze zmieniającymi się parametrami $\alpha, \beta \in \{0.5, \dots, 5\}$ (każdy z parametrów może przyjmować 10 wartości, co daje łączną liczbę 100 możliwych kombinacji par parametrów (α, β)).

6. Ocena wrażliwości procesu dotkliwości strat

Punktem wyjścia jest stworzenie teoretycznych procesów generujących dane. Przyjęto trzy warianty danych uzależnione od hipotetycznego procesu cenzurowania lub ucinania obserwacji:

- Pełny – oznacza, że rejestrowany jest pełny zakres danych. Jest to proces referencyjny, do którego będą się odwoływać pozostałe scenariusze symulacyjne.
- Ucięty – oznacza, że rejestrowane są tylko obserwacje powyżej znanego progu „c”. Brakuje natomiast wiedzy o liczbie obserwacji poniżej punktu progowego.
- Cenzurowany – oznacza, że dane powyżej progu „c” będą rejestrowane jako pełne dane, natomiast poniżej progu będą zliczane bez informacji o wielkości zmiennej losowej.

Względem trzech procesów generujących dane możliwe są trzy metody estymacji parametrów odpowiedniego ciągłego procesu dotkliwości zdefiniowane jako:

- Wariant bazowy. Dostępne są pełne dane. Standardowa metoda estymacji, tj. MLE [Klugman i in., 2008, s. 381].
- Wariant ucięty. Dostępne są dane ucięte oraz jest to uwzględnione w procesie estymacji, tj. estymowane są dane ucięte (wariant metody danej wzorem (1)).
- Wariant cenzurowany. Dostępne są dane cenzurowane, tj. rozszerzone względem uciętych o dodatkową informację o częstości występowania zjawiska poniżej progu „c”, oraz dobrana metoda estymacji (wariant metody danej wzorem (1)) uwzględnia fakt lewostronnego cenzurowania w procesie optymalizacji MLE.

Dla każdego z możliwych wariantów rozkładu $Burr3(\cdot)$ (zmiennie wartości parametrów α i β) przyjęto identyczny schemat badania symulacyjnego, które było powtórzone 10^4 razy (iteracji).

Pojedyncza iteracja w badaniu symulacyjnym dla wybranego wariantu rozkładu $Burr3(\cdot)$ zakłada, że dysponuje się hipotetyczną próbą losową 200 obserwacji (hipotetycznie cztery lata, przeciętnie po pięćdziesiąt obserwacji w każ-

dym roku). Na tej podstawie skonstruowano trzy warianty danych (bazowe – bez zmian, ucięte – informacja poniżej „c” jest tracona, cenzurowane – informacja poniżej „c” jest zliczana). W odniesieniu do wariantów zależnych od progu „c” liczba analizowanych przypadków rozszerzy się o liczbę różnych progów raportowania.

Za punkt „c”, tj. próg pełnego rejestrowania danych, przyjęto wartości kwantyli odpowiednich rzędów (ozn. $c(q)$). Wielkości kwantyli reprezentują punkty lewostronnego ucinania lub cenzurowania danych i mają one przedstawiać procentową hipotetyczną skalę utraty informacji w każdym z przypadków. Wielkości kwantyli będą się różnić w zależności od wartości wykorzystanych parametrów dla rozważanego rozkładu. Bazowanie na ustalonych rzędach kwantyli daje możliwość porównania wyników ze względu na odsetek traconej informacji. Na potrzeby badania przyjęto wielkości rzędu kwantyla odpowiednio $q \in (0,05; 0,10; \dots; 0,5)$ reprezentujące stopień cenzurowania danych rzędu od 5% do 50%.

W celu zbiorczego przedstawienia wyników przyjęto, że dla każdego z wariantów badania, tj. trzy warianty estymacji, 100 możliwych kombinacji par parametrów, wyznaczono parametry szacowanych modeli (α, β) oraz wartości miary $VaR(0,999)$, które będą wykorzystane w dalszym podsumowaniu.

W każdej iteracji, dla każdej kombinacji par parametrów (100 kombinacji) oraz każdego rzędu kwantyla (10 wartości) wyznaczone wielkości miary $VaR(0,999)$ dla wariantów estymacji cenzurowanej (ozn. VaR_C) oraz uciętej (ozn. VaR_U) zostały porównane z wartością miary $VaR_B(0,999)$ dla wariantu bazowego (tj. bez utraty informacji, oznaczenie VaR_B). Wielkość VaR_B została każdorazowo podzielona przez VaR_C oraz VaR_U . Dwie wielkości ilorazów będą w tym przypadku informowały o procentowej wielkości odchylenia wariantu bazowego od wariantów cenzurowanego oraz uciętego. Wyniki te dadzą możliwość porównania wpływu, jaki wywiera utrata informacji (reprezentowana przez rząd przyjętego kwantyla q względem parametrów rozkładu). Ponadto wielkości wspomnianych ilorazów dają możliwość wzajemnego porównania par parametrów rozkładu dotkliwości między sobą ze względu na zmienność miary $VaR(0,999)$.

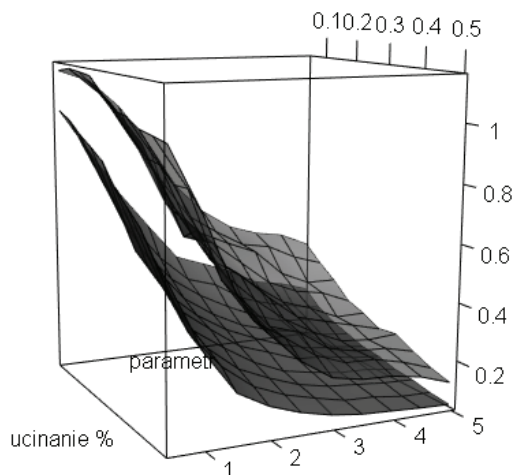
Stworzone porównania w dalszej części artykułu będą przedstawione na przestrzennych rysunkach, gdzie zawsze oś OZ będzie reprezentować pewną charakterystykę (kwantyl rzędu 1 oraz kwantyl rzędu 3) wyznaczoną dla pary zmiennych. Pary zmiennych to odpowiednio:

1. Porównanie 1 – utrata informacji (wielkość q) względem parametru α .
2. Porównanie 2 – utrata informacji (wielkość q) względem parametru β .
3. Porównanie 3 – parametr α względem parametru β .

Celem powyższych porównań jest próba odpowiedzi na pytanie, jak utrata informacji spowodowana cenzurowaniem lub ucinaniem danych będzie wpływa-

ła na zmienność wyników. W sytuacji idealnej będzie brakowało wzajemnej interakcji pomiędzy porównywanymi zmiennymi w rozważanych wariantach, co hipotetycznie będzie zobrazowane płaskimi płaszczyznami dla zastosowanych charakterystyk na poziomie osi OZ zbliżonym do wartości równej jeden w każdym z kierunków zmian wartości zmiennych z porównania. W przypadku istnienia interakcji w zależności od wartości zestawianych zmiennych z odpowiednich porównań charakter płaszczyzn będzie się zmieniał wraz z wartościami osi OX i OY. Na każdym z rysunków są widoczne dwie płaszczyzny, które przedstawiają zakres zmienności ograniczony z dołu przez wielkości kwantyli rzędu 1 oraz z góry przez wielkości kwantyli rzędu 3 dla każdej z możliwych kombinacji zmiennych.

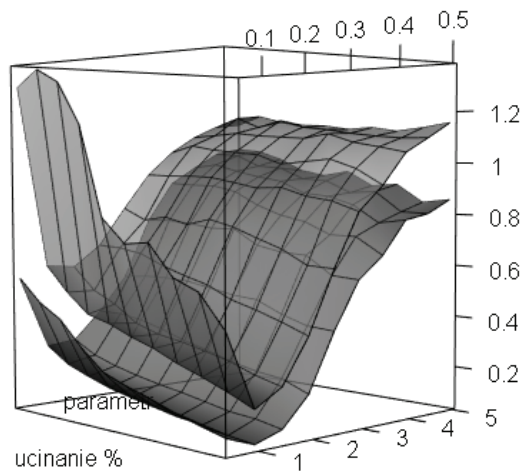
Na dwóch pierwszych rysunkach (rys. 1 oraz 2) lewa oś w płaszczyźnie poziomej jest związana z poziomem ucinania (początkowe wartości to poziom 50%, który wraz z oddalaniem się zbiega do poziomu 5%). Natomiast prawa oś (widoczne wartości od 1 do 5 w opisie osi) to wartości parametru α (rys. 1) oraz β (rys. 2) z rozkładu $Burr3(\cdot)$. W pierwszym przypadku (rys. 1) widać, że dla niewielkiego procesu ucinania i początkowych wartości parametru α wielkości z osi OZ, tj. odchylenia od wariantu bazowego są niewielkie.



Rys. 1. Wariant ucięty. Iloraz VaR_B/VaR_U . Porównanie q/α

Wzrost procesu ucinania oraz wzrost wartości parametru α powoduje, że potencjalne różnice pomiędzy wariantem bazowym a uciętym znacząco się zmieniają. W skrajnym przypadku, tj. dla 50% utraty informacji, wielkości miary VaR wyznaczone z wariantu uciętego są znacznie (kilkukrotnie) większe od wariantu bazowego, o czym świadczą wartości osi OZ bliskie wartości zero.

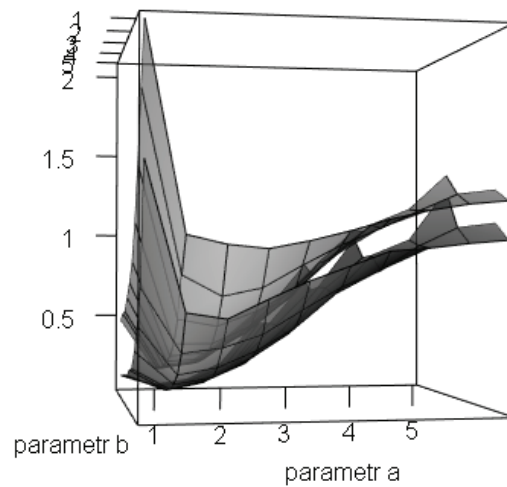
W przeciwieństwie do wyników z rys. 1, porównując wielkości q do kolejnego parametru rozkładu, tj. wielkości β , można zauważyć odmienny charakter powierzchni, jaka powstaje przez naniesienie ograniczeń z wyznaczonych płaszczyzn. Tutaj główna zmiana jest widoczna wzdłuż prawej poziomej osi, tj. wraz ze zmianą parametru β . Ruch względem postępującego procesu ucinania jest największy dla wartości parametru $\beta < 2,5$, dalej niewielkie zmiany wzdłuż osi zmiany wielkości q świadczą o niewielkim już wpływie procesu ucinania przy takich wielkościach parametrów.



Rys. 2. Wariant ucięty. Iloraz VaR_B/VaR_U . Porównanie q/β

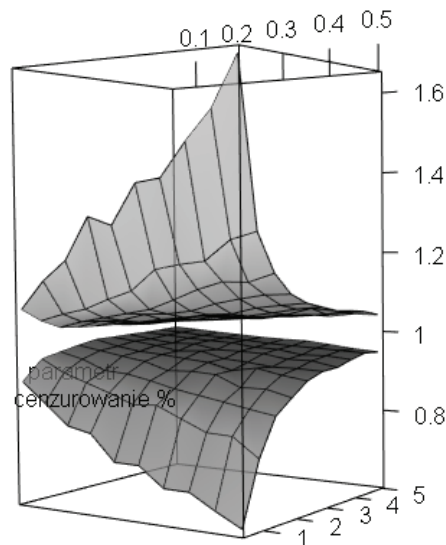
Zmienność wielkości miary VaR (wyznaczonych tylko dla procesu dotkliwości) ma niejednorodny charakter, który wraz ze wzrostem parametru β stabilizuje się w granicach wartości osi OZ równych 1 (świadczy to o zbieżności wyników w tym zakresie dziedziny w odniesieniu do wariantu bazowego).

Do podobnych wniosków można dojść analizując wykres z rys. 3, gdzie iloraz wielkości VaR dla wzajemnej relacji dwóch parametrów rozkładu $Burr3(\cdot)$ stabilizuje się dla wyższych wartości.



Rys. 3. Wariant ucięty. Iloraz VaR_B/VaR_U . Porównanie α/β

W przypadku kolejnych wykresów (rys. od 4 do 6) dotyczących wpływu cenzurowania informacji widać odmienne interakcje pomiędzy analizowanymi zmiennymi z poszczególnych wariantów.

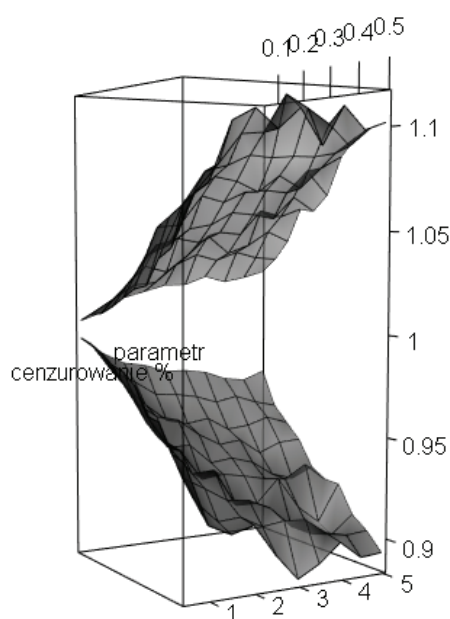


Rys. 4. Wariant cenzurowany. Iloraz VaR_B/VaR_C . Porównanie q/α

Na rysunku 4, gdzie zostały porównane wielkości q (tu utrata danych jest częściowa, a informacja o cenzurowaniu danych jest wykorzystywana w procesie estymacji) z parametrem α analizowanego rozkładu, można zauważyć naj-

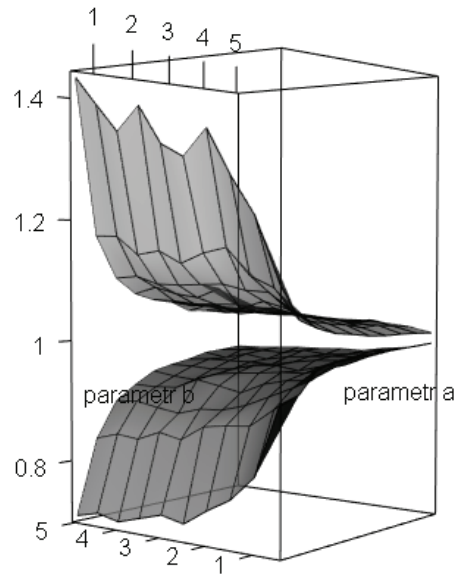
większe różnice według skali osi OZ, dla początkowych wartości parametru. Dla tych wartości α , które są mniejsze od wartości około 2,5, widać również największą interakcję z procesem cenzurowania. W przypadku pozostałej dziedziny α wpływ cenzurowania nie miał większego znaczenia.

W przypadku kolejnego zestawienia widocznego na rys. 5 można zauważyć, że cenzurowanie odgrywało tam największą rolę w rozrzucie wielkości miary VaR_B w odniesieniu do VaR_C . Należy mieć jednak na względzie, że skala osi OZ jest nieznaczna (zmiennosc od 0,9 do 1,1), stąd w praktyce wrażliwość szacowanej wielkości $VaR(0,999)$, nawet przy dużej utracie informacji (tj. cenzurowanie na poziomie 50%), jest niewielka.



Rys. 5. Wariant cenzurowany. Iloraz VaR_B/VaR_C . Porównanie q/β

Porównując również wzajemne interakcje obu parametrów, jakie są prezentowane na rys. 6, można zauważyć, że odpowiedzialny za największą zmienność wyników jest parametr β . Jednak wraz z większymi wartościami, tj. od około wielkości 2,5, znaczenie cenzurowania nawet dla dużych wielkości ma niewielką siłę oddziaływania.



Rys. 6. Wariant cenzurowany. Iloraz VaR_B/VaR_C . Porównanie α/β

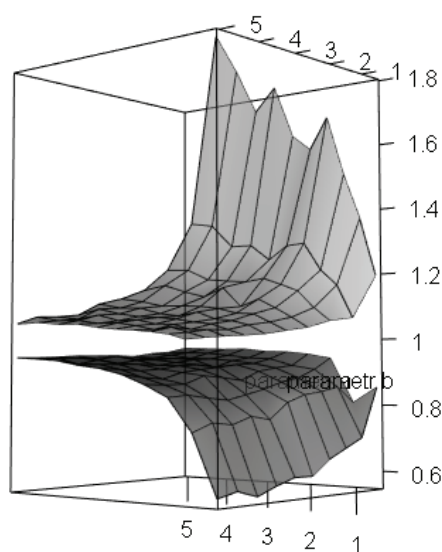
Zaprezentowane wyniki pierwszego badania symulacyjnego pokazują, że wariant estymacji uciętej, tj. sytuacji, w której tracona jest bezpowrotnie część istotnych informacji, będzie powodował dużą zmienność i wrażliwość wyników uzależnioną od specyficznych wielkości parametrów rozkładu, jak również od wielkości utraty informacji.

7. Wpływ cenzurowania i ucinania obserwacji na zagregowany rozkład strat

W tej części badania symulacyjnego wyznaczono wielkość miary $VaR(0,999)$ ze zagregowanego rozkładu strat. W celu zaprezentowania wpływu ucinania i cenzurowania informacji przyjęto, jak poprzednio, trzy metody estymacji w zależności od wariantu danych. Z uwagi na fakt, że metoda LDA korzysta z dwóch procesów cząstkowych, tj. procesu częstości (dyskretny rozkład prawdopodobieństwa) oraz procesu dotkliwości (proces ciągły), to w przypadku ucinania obserwacji tracony będzie odpowiedni procent obserwacji. Przyjęto w tym przypadku, że procesem wyjściowym będzie jednorodny proces Poissona. Wstępne założenia co do średniej liczby zdarzeń operacyjnych przyjęto na poziomie 50 obserwacji w ciągu roku, co będzie równe parametrowi λ w rozkładzie Poissona ze stałą intensywnością.

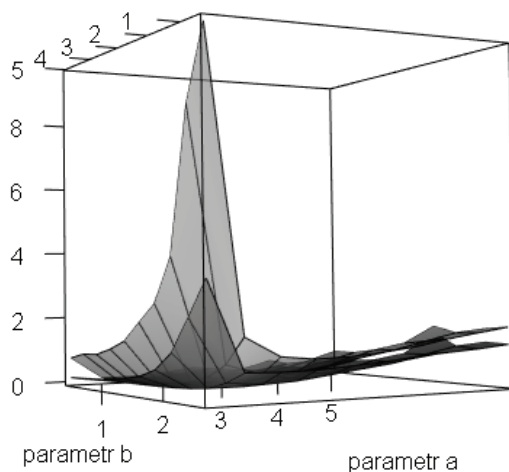
W przypadku rozwiązania bazowego (pełna wiedza o zdarzeniach) oraz rozwiązania cenzurowania danych poniżej określonego progu (zliczanie określonych informacji) intensywność zdarzeń pozostanie bez zmian. Natomiast w przypadku ucinania obserwacji poziom q (reprezentujący procentowy poziom utraty informacji) będzie informował o zmianie częstości zdarzeń. Przyjęto, że pierwotna częstość $E(N) = \lambda$ będzie zastąpiona wielkością $(1 - q)\lambda$. Postępowanie takie wynika z własności rozkładu Poissona, jakimi są twierdzenie o dodawaniu oraz twierdzenia Rajkowa [por. Fisz, 1969, s. 103 i 157; Shevchenko, 2011, s. 181].

Zbiorcze wyniki zaprezentowano na poniższych rysunkach (rys. 7 oraz 8). Zestawiono tam w identyczny sposób, jak w przypadku wcześniejszych rysunków (rys. 6 oraz 3), parametry rozkładu $Burr3(\cdot)$ względem ilorazów wielkości $VaR_B^Z(0,999)/VaR_C^Z(0,999)$ oraz $VaR_B^Z(0,999)/VaR_U^Z(0,999)$. Widać pewne podobieństwo w kształtach dla odpowiednich przypadków, tj. porównania względem wariantu cenzurowanego (rys. 7) oraz uciętego (rys. 8). Są to znaczące wartości na osi OZ informujące o skali potencjalnych różnic pomiędzy odpowiednimi miarami ryzyka. W przypadku zagregowanych rozkładów strat, tj. wyników zaprezentowanych zbiorczo na rys. 7 oraz 8, widać większe rozbieżności niż w przypadkach pojedynczych rozkładów dotkliwości.



Rys. 7. Wariant cenzurowany. Iloraz VaR_B/VaR_C . Porównanie α/β

W przypadku wariantu cenzurowania efekt wynikający z utraty informacji nie jest duży i jedynie w przypadku początkowych wielkości parametrów z rozkładu dotkliwości widać prawie dwukrotne różnice w wielkościach miar ryzyka wyznaczonych ze zagregowanych rozkładów bazowego oraz cenzurowanego.



Rys. 8. Warianc cenzurowany. Iloraz VaR_B/VaR_C . Porównanie α/β

W przypadku wariantu ucinania informacji skala potencjalnych różnic jest znacząca dla początkowych wartości parametrów rozkładu dotkliwości strat, widać ponad dziesięciokrotne różnice, które wraz ze wzrostem wielkości tych parametrów zaczynają być zbieżne z wariantem bazowym.

Drugi etap badań pokazał, że charakter zależności pomiędzy tymi samymi rodzajami różnic (różne ilorazy wielkości $VaR(\cdot)$) w wielkościach miar ryzyka dla różnych wariantów jest zbieżny dla rozkładów zagregowanych i pojedynczych rozkładów dotkliwości strat. Ewentualna znacząca różnica wynika jedynie z utraty informacji dotyczącej częstości. Niemniej jednak informacja o cenzurowaniu pozwala na znaczące zmniejszenie zmienności miary ryzyka wynikającej z rozkładu dotkliwości strat. W przypadku danych uciętych skala potencjalnych różnic w wielkościach szacowanego ryzyka jest odmienna dla różnych wartości parametrów rozkładów dotkliwości. Należy podkreślić, że w każdym z rozważanych scenariuszy skala różnic wynikająca z procesu ucinania ma tendencje do stabilizowania się i zbiegania do wariantu bazowego wraz ze wzrostem wartości parametru rozkładu dotkliwości, jakim jest w tym przypadku rozkład $Burr3(\cdot)$.

Podsumowanie

Analizując zaprezentowane wyniki z dwóch etapów badania symulacyjnego, należy stwierdzić, że proces ucinania danych, który jest stosowany powszechnie w ryzyku operacyjnym, może być obciążony dużą zmiennością wyników. Potencjalne różnice w przypadku utraty nie tylko danych dla rozkładu dotkliwości,

ale także informacji o częstotliwości zdarzeń mogą w niektórych przypadkach powodować skrajne niedoszacowanie wielkości ryzyka ze zagregowanego rozkładu strat. Niedoszacowanie dotyczy sytuacji, w których prezentowane ilorazy miar wielkości VaR jako wartości na osiach OZ na wszystkich prezentowanych rysunkach przyjmują wartości większe od jeden.

W przypadku wykorzystania cenzurowania jako procesu rejestracji danych powoduje to, że cała dziedzina częstości jest we właściwy sposób reprezentowana. Ponadto utrata informacji w procesie dotkliwości nie ma tak drastycznego oddziaływania na wielkości szacowanego ryzyka.

Korzyści wynikające z cenzurowania danych mogą być nie tylko wykorzystane w przypadku prostego szacowania wielkości ryzyka operacyjnego. Po pierwsze cenzurowanie daje możliwość cyklicznego raportowania strat przez poszczególne komórki banku, co w znaczący sposób nie zmieni ich efektywności (raportowanie może być w podokresach, np. tygodniowych miesięcznych – w zależności od tego, w jakim układzie są szacowane parametry rozkładów częstości). Po drugie cenzurowanie w metodzie LDA może znaleźć zastosowanie jako pewnego rodzaju benchmark, jak również może być wykorzystane w przypadku realizacji testów warunków skrajnych czy ogólnym badaniu odporności stosowanych w praktyce rozwiązań na niekorzystne zmiany wynikające z procesów cząstkowych. Po trzecie cenzurowanie daje realne możliwości łączenia danych pochodzących z różnych źródeł, dla których zazwyczaj istnieją różne progi raportowania strat operacyjnych. Wynika to bezpośrednio ze sposobu cenzurowania typu II, w przypadku którego progresywne, tj. narastające progi cenzurowania nie ograniczają w znaczący sposób możliwości estymacji parametrów procesu dotkliwości strat.

Zaprezentowana w artykule tematyka związana z pomiarem wielkości ryzyka operacyjnego wskazuje, że niedoszacowanie stopnia narażenia na ryzyko operacyjne może być spowodowane błędnym postrzeganiem procesu generującego dane. Zmiana koncepcji zbierania i wykorzystania danych, która bazuje na procesach lewostronnego cenzurowania strat, może stanowić częściowe rozwiązanie tych problemów i nie powinna być pomijana w praktyce bankowej, w szczególności przy szacowaniu kapitału ekonomicznego.

Literatura

BIS (2010, October 28), *Recognising the Risk-mitigating Impact of Insurance in Operational Risk Modelling*, <http://www.bis.org/publ/bcbs181.htm>.

BIS (2011, June 30), *Principles for the Sound Management of Operational Risk – final document*, <http://www.bis.org/publ/bcbs195.htm>.

- BIS (2013, January 9), *Principles for Effective Risk Data Aggregation and Risk Reporting*, <http://www.bis.org/publ/bcbs239.htm>.
- BIS (2014, October 6), *Operational Risk – Revisions to the Simpler Approaches- Consultative Document*, <http://www.bis.org/publ/bcbs291.htm>.
- BIS (2016, March 4), *Standardised Measurement Approach for Operational Risk – consultative document*, <http://www.bis.org/publ/d355.htm>.
- Chernobai A.S., Rachev S.T. & Fabozzi F.J. (2007), *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*, John Wiley & Sons.
- Feroze N., Aslam M. and Sindhu T.N. (2015), *On Maximum Likelihood Estimation for Left Censored Burr Type III Distribution*, "Pakistan Journal of Statistics and Operation Research", Vol. 11, No. 4, s. 497-512.
- Fisz M. (1969), *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, Warszawa.
- Klein J.P., Moeschberger M.L. (2003), *Survival Analysis: Techniques for Censored and Truncated Data*, Springer.
- Klugman S.A., Panjer H.H. & Willmot G.E. (2008), *Loss Models: From Data to Decisions*, Wiley.
- Millard S.P. (2013), *EnvStats*, Springer, New York.
- PWC (2015 December), *Operational Risk: The End of Internal Modelling?* <https://www.pwc.com/gx/en/financial-services/pdf/fs-operational-risk-modelling.pdf>.
- Shevchenko P.V. (2011), *Modelling Operational Risk Using Bayesian Inference*, Springer, Berlin-Heidelberg.
- Szkutnik T., Basiaga K. (2013), *The Application of Generalized Pareto Distribution and Copula Functions in the Issue of Operational Risk*, "Econometrics", 39(1), s. 133-142.

INFLUENCE OF LEFT CENSORING OF SEVERITY OF LOSSES ON ESTIMATING OPERATIONAL RISK AT THE EXAMPLE OF BURR TYPE III DISTRIBUTION

Summary: The article is addressed to the problem of using various types of reporting in regard to estimating the operational risk with LDA method. Advanced Measurement Approach of estimation in the latest solutions of the Basel Committee on Banking Supervision will not probably be used anymore for defining the size of regulatory capital, however, its usefulness may remain one of elements of risk management system in financial institution. The aim of the article is to examine an impact of loss of information about operational losses on stability of risk estimates performed by LDA method. In that context, solutions which use left censored distributions that are proposed in the article, may be an interesting alternative in practical solutions. The article presents the problem of left censored Burr type III distribution, which may be foundation to determine severity of losses distribution. Considered concept that is basing on censored sample was compared to truncation which is very common solution in practice. Simulations has showed

that in case of censoring of data we can expect more stable results which will be closer to situation of full knowledge about happening processes on which operational risk is basing. Estimation of parameters on the basis of censored data used commonly in other research areas may be one of directions of development also in case of operational risk in financial institutions.

Keywords: operational risk, LDA, left censoring, Burr distribution.