



Maciej Pichura

Uniwersytet Ekonomiczny w Katowicach
Wydział Ekonomii
Katedra Metod Statystyczno-Matematycznych w Ekonomii
maciej.pichura@gmail.com

MOŻLIWOŚCI ZASTOSOWANIA METOD WTÓRNEGO PRÓBKOWANIA DO WERYFIKACJI EFEKTYWNOŚCI INWESTYCJI PORTFELOWYCH

Streszczenie: Artykuł przedstawia możliwość adaptacji jednej z metod wtórnego próbkowania – powtarzanej K-składowej walidacji krzyżowej, w odniesieniu do oceny efektywności inwestycyjnej strategii portfelowych. Metoda ta pozwala na wnioskowanie statystyczne w sytuacjach, gdy nie jest znana postać analityczna rozkładu badanej cechy lub postać ta nie pozwala na skuteczne zastosowanie dostępnych metod parametrycznych. Badanie empiryczne, które przeprowadzono w celu potwierdzenia stawianej hipotezy nie wskazuje na fakt, iż należałoby jej zaprzeczyć. Zostało ono wykonane na dużym zbiorze danych historycznych pochodzących z amerykańskiego rynku akcji. Dodatkowo w artykule zaprezentowano również sposoby wyznaczania błędów prognozy mierników efektywności inwestycyjnej podczas stosowania metod wtórnego próbkowania.

Słowa kluczowe: K-składowa walidacja krzyżowa, wtórne próbkowanie, metody portfelowe, efektywność inwestycji.

Wprowadzenie

Obszar nauki zajmujący się inwestycjami kapitałowymi jest bardzo rozległy i charakteryzuje się mnogością metod i strategii inwestycyjnych. Większość z nich charakteryzuje trudność weryfikacji statystycznej ich rezultatów, która spowodowana jest specyficznymi cechami badanych szeregów statystycznych. Są to szeregi czasowe stóp zwrotu, które często nie mają znanej postaci analitycznej rozkładu, a także wykazują brak stacjonarności i homoskedastyczności. Dodatkowo szeregi te w wielu wypadkach mają stosunkowo niewielką liczebność obserwacji

historycznych, co spowodowane jest różnorodnym stopniem rozwoju rynków kapitałowych w poszczególnych krajach.

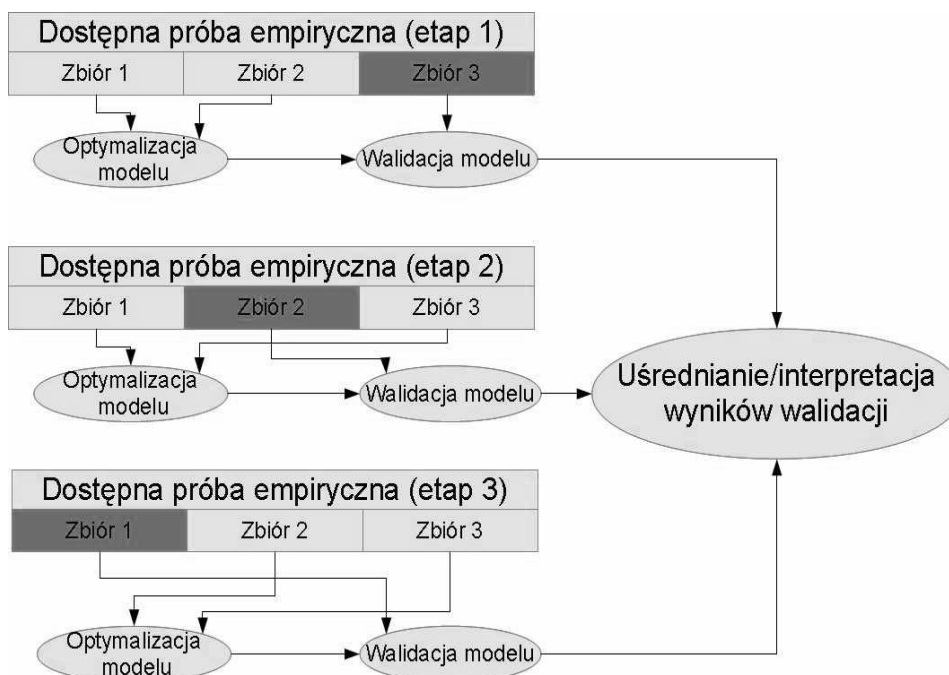
Aby problem ten przezwyciężyć, można zastosować metody wtórnego próbkowania, które pozwalają na uzyskanie znacznie większej liczby obserwacji, co może również pomóc w budowaniu rozkładu prawdopodobieństwa, a co za tym idzie wnioskowaniu statystycznym. W sensie ogólnym wtórne próbkowanie (*re-sampling*) polega na wielokrotnym powtarzaniu procedur, które dają w wyniku szeregi statystyczne na próbach osiąganych z próby pierwotnej poprzez jej dzielenie, powielanie lub losowanie z niej podprób [Efron, 1982].

Głównym celem tego artykułu jest przedstawienie jednej z możliwych metod wtórnego próbkowania – powtarzanej walidacji krzyżowej – oraz jej zastosowania w ocenie efektywności inwestycji portfelowych. Hipoteza badawcza, która zostanie tutaj postawiona polega na założeniu, że procedury wtórnego próbkowania pozwalają na stosowanie metod wnioskowania statystycznego w weryfikacji rezultatów inwestycji bez przyjmowania założeń dotyczących postaci analitycznej rozkładów ich stóp zwrotu.

1. Metody wtórnego próbkowania – walidacja krzyżowa

Procedury wtórnego próbkowania znajdują podstawowe zastosowanie w problemach związanych z wnioskowaniem statystycznym. Są to metody z szeroko opisywanej w literaturze grupy metod *bootstrap*, natomiast istnieje jeszcze jedna istotna dziedzina, w której są aplikowane. Jest to obszar problemów regresyjnych, w którym trudności przysparza często prawidłowa weryfikacja dokładności prognozy modelu. W celu lepszej oceny trafności prognoz już w latach 30. XX w. zaproponowano metody weryfikacji modeli regresji poprzez rozdział dostępnych danych na dwa rozłączne zbiory – uczący i testowy. Polegała ona na optymalizacji modelu na zbiorze uczącym i jego walidacji na zbiorze testowym. Metoda była cały czas rozwijana, a kilka dekad później jej modyfikacje i rozszerzenia zostały zakwalifikowane do grupy metod walidacji krzyżowej (*cross validation*) [Liu, Ozs, red., 2009, s. 532-538]. Jedną z najpopularniejszych metod z tej grupy jest tak zwana K-składowa walidacja krzyżowa (*K-fold cross validation*). Polega ona na podzieleniu dostępnego zbioru danych na K podzbiorów o równej (lub bardzo zbliżonej) liczebności i wykonaniu K optymalizacji analizowanego modelu. W każdej iteracji wybiera się jeden zbiór testowy, na którym waliduje się model. Pozostałe K-1 zbiorów łączy się w jeden i traktuje jako zbiór uczący. W ten sposób każdy z wyodrębnionych zbiorów służy dokładnie jeden raz jako

zbiór walidacyjny oraz jest $K-1$ razy składową zbioru uczącego. Rezultaty uzyskane w iteracjach można uśrednić lub przetwarzać w inny wybrany sposób, który jest odpowiedni dla przyjętego modelu. Ilustracja graficzna sposobu przeprowadzania 3-składowej walidacji krzyżowej została przedstawiona na rysunku 1.



Rys. 1. Przykład procedury 3-składowej walidacji krzyżowej

Najbardziej rozwiniętą odmianą metody walidacji krzyżowej jest K -składowa powtarzana walidacja krzyżowa (*K-fold repeated cross validation*). Polega ona na wielokrotnym powtarzaniu procedury K -składowej walidacji krzyżowej. W każdym powtórzeniu podział dostępnej próby empirycznej na podzbiory jest wykonywany w sposób losowy. Zachowuje się przy tej operacji pewne zasady co do liczebności wszystkich podzbiorów, aby w wypadku każdego zbioru testowego pozwalała na wykonanie wiarygodnej walidacji modelu. Stosując takie podejście należy również zwrócić uwagę na fakt, iż zbiory uczące oraz zbiory testowe z różnych powtórzeń procedury mogą mieć elementy wspólne. Mając na uwadze powyższe, w procedurze K -składowej powtarzanej walidacji krzyżowej można odnaleźć wyraźne analogie do metod *bootstrap*.

Spośród kluczowych zastosowań grupy metod walidacji krzyżowej można wyróżnić podstawowe, którym jest narzędzie weryfikacji jakości prognostycznej

modelu. W tym wypadku analizuje się bowiem najczęściej uśredniony błąd prognozy [Fushiki, 2011, s. 137-146; Bunke, Droge, 1984, s. 1400-1424]. Jako parametr ten w tradycyjnych problemach regresji najczęściej zastosowanie znajduje średniokwadratowy błąd prognozy. Dla pojedynczego etapu procedury K-składowej walidacji krzyżowej można go wyrazić jako [Sobczyk, 2000]:

$$se = \sqrt{\frac{1}{N_K} \sum_{i=1}^{N_K} (y_i - y_i^*)^2}, \quad (1)$$

gdzie:

- se – średniokwadratowy błąd prognozy,
- N_K – liczebność danego zbioru testowego,
- y_i – rzeczywista wartość o indeksie i zaobserwowana w zbiorze testowym,
- y_i^* – prognozowana wartość o indeksie i .

Dla pojedynczego powtórzenia procedury K-składowej walidacji krzyżowej można zatem wyznaczyć wartość przeciętną średniokwadratowego błędu prognozy i wyrazić ją w następujący sposób [Tibshirani, Tibshirani, 2009, s. 822-829]:

$$SE_{CV} = \frac{1}{K} \sum_{i=1}^K se_i, \quad (2)$$

gdzie:

- SE_{CV} – przeciętny średniokwadratowy błąd prognozy metody K-składowej walidacji krzyżowej,
- K – liczba etapów walidacji,
- se_i – średniokwadratowy błąd prognozy dla etapu walidacji oznaczonego indeksem i .

W związku z tym, że przedstawiony we wzorze (1) przeciętny błąd prognozy jest estymatorem rzeczywistego oczekiwanego błędu należy podjąć próbę wyznaczenia jego obciążenia. Jeden z zaproponowanych w literaturze sposobów określenia obciążenia estymatora oczekiwanego błędu prognozy ma następującą postać [Tibshirani, Tibshirani, 2009, s. 822-829]:

$$b(SE_{CV}) = \frac{1}{K} \sum_{i=1}^K (se_i - se_{\min}), \quad (3)$$

gdzie:

$b(SE_{CV})$ – obciążenie błędu prognozy metody K-składowej walidacji krzyżowej,
 se_{\min} – minimalna wartość błędu prognozy spośród uzyskanych we wszystkich etapach procedury walidacji,

Pozostałe oznaczenia jak we wzorze (2).

Weryfikacja trafności prognoz opisana w powyższy sposób może również znaleźć zastosowanie w metodzie K-składowej powtarzanej walidacji krzyżowej. Konieczny do uwzględnienia w takim wypadku jest fakt, iż wartości błędu prognozy oraz jego obciążenia w pojedynczym powtórzeniu procedury muszą zostać dodatkowo uśrednione po wykonaniu zadanej liczby ich iteracji.

W wypadku weryfikacji efektywności metody lub strategii inwestycyjnej możliwa jest adaptacja procedury K-składowej walidacji krzyżowej. Jako model należy w takiej adaptacji przyjąć analityczną postać metody inwestycyjnej, natomiast jako zmienną objaśnianą – miarę efektywności stosowaną do oceny rezultatów (wyrażona na przykład jako φ), którą oblicza się na podstawie szeregu stóp zwrotu, który generuje metoda inwestycyjna. Dla powyższych założeń można określić:

$$SE_{CV\varphi} = \sqrt{\frac{1}{K} \sum_{i=j=1}^K (\hat{\varphi}_j - \hat{\varphi}_i)^2}, \quad (4)$$

gdzie:

$SE_{CV\varphi}$ – średniokwadratowy błąd prognozy miernika efektywności w procedurze K-składowej walidacji krzyżowej,

K – liczba podprób procedury K-składowej walidacji krzyżowej,

$\hat{\varphi}_j$ – wartość estymatora miernika efektywności wynikająca z zastosowania metody/strategii inwestycyjnej na podpróbie testowej o indeksie j ,

$\hat{\varphi}_i$ – wartość estymatora miernika efektywności wynikająca z zastosowania metody/strategii inwestycyjnej na podpróbie uczącej o indeksie i .

Interpretacja i ocena średniokwadratowego błędu prognozy przysparza w niektórych przypadkach dużych trudności. Oczywiście jest bowiem, że wartość błędu średniokwadratowego powinna być jak najmniejsza, natomiast trudności nastęrcza określenie, które wartości są wystarczająco niskie, żeby uznać je za zadowalające. W wypadku metody powtarzanej walidacji krzyżowej najlepszym rozwiązaniem tego problemu jest zastosowanie względnego błędu prognozy, który można określić wzorem:

$$\text{VSE}_{CV\varphi} = \sqrt{\frac{1}{K} \sum_{i=1}^K \left(\frac{\hat{\varphi}_j - \hat{\varphi}_i}{\hat{\varphi}_j} \right)^2} \quad (5)$$

gdzie:

$\text{VSE}_{CV\varphi}$ – względny średniokwadratowy błąd prognozy miernika efektywności w procedurze K-składowej walidacji krzyżowej,

Pozostałe oznaczenia jak we wzorze (4).

Parametry opisane wzorami (4) i (5) w metodzie K-składowej powtarzanej walidacji krzyżowej mogą znaleźć wielorakie zastosowania. Najprostszym z nich jest uśrednienie ich wartości, natomiast możliwe jest również uzyskanie wartości ich miar pozycyjnych, a także percentylowego przedziału ufności, który można wyrazić wzorem:

$$P \left(\hat{\theta} - q_{\frac{\alpha}{2}}^* < \theta < \hat{\theta} + q_{1-\frac{\alpha}{2}}^* \right) = 1 - \alpha, \quad (6)$$

gdzie:

q_{α}^* – percentyl rzędu $\alpha/100$ rozkładu empirycznego $\hat{\theta}$ wygenerowanego poprzez kolejne iteracje procedury K-składowej walidacji krzyżowej,

$\hat{\theta}$ – estymowany parametr ($\text{SE}_{CV\varphi}$ lub $\text{VSE}_{CV\varphi}$).

Powyższy wzór znajduje zastosowanie jedynie w wypadkach, gdy uzyskany w procedurze rozkład empiryczny jest symetryczny lub zbliżony do symetrycznego.

2. Wyniki badań empirycznych

W związku z przeprowadzanymi w artykule rozważaniami zostały wykonane empiryczne badania mające na celu ułatwić wnioskowanie w podejmowanym tutaj zagadnieniu. Ich głównym założeniem jest przeprowadzenie procedury 3-składowej powtarzanej walidacji krzyżowej w odniesieniu do oceny efektywności portfela inwestycyjnego. Badanie w założeniu ma wykazać, iż do weryfikacji statystycznej tejsze efektywności nie jest konieczne domniemanie postaci analitycznej rozkładu stóp zwrotu. Zastosowana metoda pozwala dodatkowo na ogólne odejście od konieczności czynienia założeń ze względu na postać analityczną analizowanego rozkładu.

Ze względu na najlepszą jakość danych empirycznych w aspekcie możliwości zastosowań metod portfelowych, których efektywność jest tutaj analizowana, do badania wybrano szeregi czasowe tygodniowych stóp zwrotu z wybranych akcji wchodzących w skład amerykańskiego indeksu S&P500. Jako kryteria doboru spółek wybrane zostały jak najwyższa wartość kapitalizacji, najwyższa płynność oraz długa dostępna historia notowań. Lista spółek oraz okres dostępnych notowań historycznych zostały przedstawione w tabeli 1.

Tabela 1. Spółki dobrane do tworzenia portfeli inwestycyjnych wraz z okresami dostępnych notowań historycznych

Nazwa spółki	Skrót nazwy	Okres dostępnych notowań	Liczoność szeregu stóp zwrotu
American Express Company	AXP	od 14.01.1972 do 30.11.2013	15 297
Bank of America	BAC	od 5.06.1986 do 30.11.2013	10 041
Chevron Corp.	CVX	od 9.01.1970 do 30.11.2013	16 032
Coca-Cola Company	KO	od 9.01.1970 do 30.11.2013	16 032
Exxon Mobil Corp.	XOM	od 9.01.1970 do 30.11.2013	16 032
General Electric Company	GE	od 9.01.1962 do 30.11.2013	18 954
IBM Corporation	IBM	od 9.01.1962 do 30.11.2013	18 954
Johnson & Johnson	JNJ	od 9.01.1970 do 30.11.2013	16 032
Pfizer Inc.	PFE	od 11.01.1982 do 30.11.2013	11 647
Procter & Gamble Company	PG	od 9.01.1970 do 30.11.2013	16 032

Jak można zauważyć, wybrane do badania szeregi stóp zwrotu są bardzo liczne, co wskazuje, iż wybrane próby empiryczne są wystarczająco reprezentatywne. Co najważniejsze, wybrane do analizy rozkłady stóp zwrotu pochodzą z najbardziej rozwiniętego i płynnego rynku akcji na świecie – amerykańskich giełd NYSE i NASDAQ. Ze względu na najbardziej pełną możliwość dywersyfikacji portfela okres badawczy zawężono do przedziału od 9.01.1970 do 30.11.2013. Za cały ten okres są bowiem dostępne notowania co najmniej siedmiu z dziesięciu wybranych spółek.

Jeśli chodzi o metodę doboru optymalnego składu portfela inwestycyjnego, która została wybrana do przeprowadzenia badania, to zastosowanie znajduje tutaj klasyczna metoda Markowitza [1952]. Jest to metoda, której kryterium optymalizacji jest osiągnięcie minimalnej wariancji stóp zwrotu z portfela przy zachowaniu założonego minimalnego poziomu stopy zwrotu. Ze względu na jej rozpowszechnienie nie zostanie ona bardziej szczegółowo przedstawiona.

Mierniki zastosowane do oceny efektywności, które wybrano do analizy, także należą do najpowszechniej znanych i stosowanych w praktyce rynkowej. Są nimi średnia stopa zwrotu z portfela oraz współczynnik Sharpe'a [Sharpe, 1994]. Również w tym wypadku nie zostaną one poddane bardziej szczegółowej prezentacji.

Uwaga zostanie natomiast skierowana na jak najbardziej wyczerpujące przedstawienie założeń i poszczególnych etapów procedury 3-składowej powtarzanej walidacji krzyżowej, a także sposobu analizy jej wyników. W pierwszej kolejności zostanie omówiony sposób wyodrębniania podprób z dostępnej próby empirycznej. Do pierwszego powtórzenia procedury walidacji krzyżowej cała dostępna próba została podzielona na trzy zbiory o możliwie równej liczebności. Zbiory te zawsze są wzajemnie wykluczające się. W kolejnych powtórzeniach podział na trzy równe podpróby został wykonywany losowo. Losowanie elementów do tych podprób nastąpiło z dostępnej próby empirycznej z założeniem, że wylosowanie każdego elementu jest równo prawdopodobne oraz że nie następuje zwracanie elementów już wylosowanych. Dla każdego powtórzenia procedury walidacji krzyżowej została obliczona wartość średniej stopy zwrotu z portfela oraz współczynnika Sharpe'a portfela zarówno dla zbioru uczącego, jak i zbioru testowego. W pierwszej części artykułu miernik efektywności został oznaczony w ogólny sposób jako φ i stosując to oznaczenie przedstawione zostały wzory (4) i (5) służące do wyznaczenia błędów prognozy tego miernika. Wzory te zostały zastosowane dla wspomnianych średniej stopy zwrotu i współczynnika Sharpe'a w każdym powtórzeniu procedury. Po wykonaniu pięciuset powtórzeń procedury walidacji krzyżowej i uzyskaniu tylu samo średniokwadratowych błędów prognozy wybranych mierników efektywności inwestycyjnej przy pomocy wzoru (6) został oszacowany przedział ufności dla wyznaczonych błędów względnych ($VSE_{CV\varphi}$). Jako dodatkowe parametry pomocne w weryfikacji efektywności metod portfelowych zostały zastosowane średnia, mediana oraz pozostałe kwartyle uzyskanych błędów prognozy. Uzyskane w przedstawiony sposób wyniki zostały przedstawione w tabeli 2.

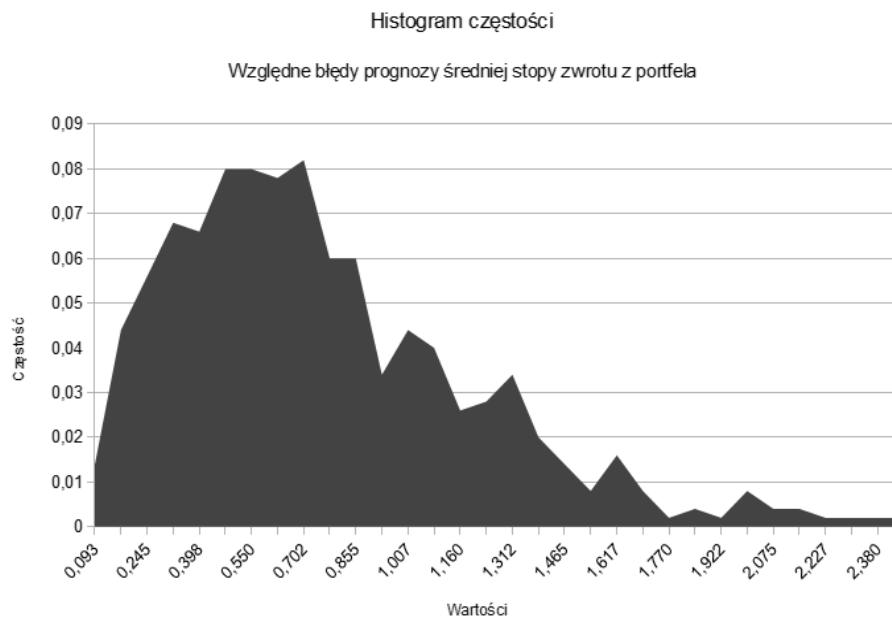
Tabela 2. Wybrane własności i parametry statystyczne względnych błędów prognozy uzyskanych w wyniku wykonania procedury 3-składowej powtarzanej walidacji krzyżowej

Własność/parametr	Wartość dla średniej stopy zwrotu	Wartość dla współczynnika Sharpe'a
1	2	3
Przedział ufności ($\alpha = 0,1$)	(0,317;1,232)	(0,264;1,102)
Średnia	0,403	0,351

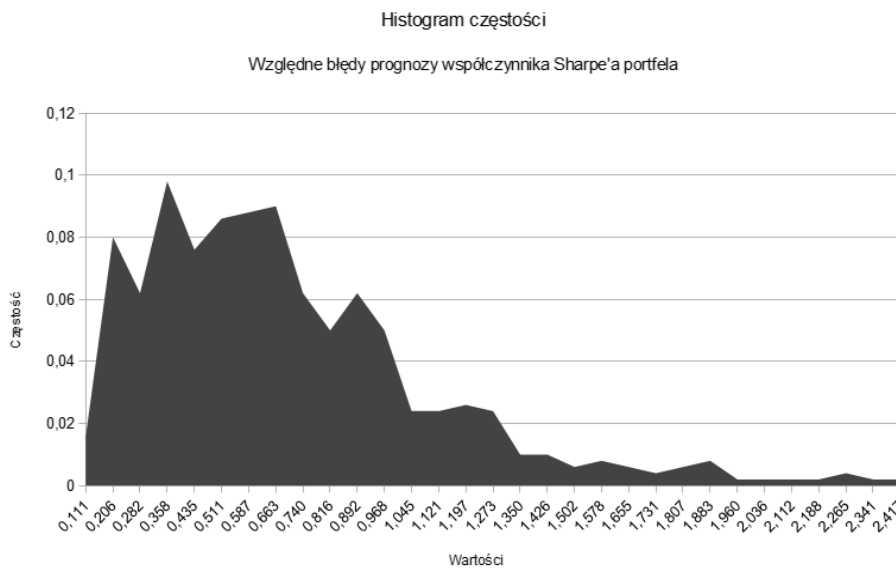
cd. tabeli 2

1	2	3
Mediana	0,346	0,295
Odchylenie standardowe	0,285	0,274
Współczynnik skośności	2,369	3,349
Kurtoza	13,213	25,195
Kwartył górny	0,218	0,175
Kwartył dolny	0,529	0,453
Percentyl 5	0,087	0,087
Percentyl 95	0,829	0,752

Średnie wartości względnego błędu prognozy oraz jego mediana dla obydwu wybranych mierników efektywności inwestycyjnej są względnie wysokie, co może wskazywać na fakt, iż na wybranej próbie empirycznej zastosowana metoda portfelową nie pozwala na uzyskiwanie powtarzalnych wyników inwestycyjnych. Wysokie odchylenie standardowe badanych błędów, które wynosi ponad 70% średniej, może wskazywać na znaczne rozproszenie uzyskanych rozkładów empirycznych. Z drugiej strony kurtoza jest bardzo wysoka, kwartyle górny i dolny wydają się nie być silnie oddalone od mediany czy też średniej. Wartość percentyla 5, a w szczególności 95, wykazuje natomiast już znaczne oddalenie od środka rozkładu. Współczynnik asymetrii obydwóch uzyskanych rozkładów empirycznych jest silnie dodatni. Biorąc pod uwagę te dane, można stwierdzić, że rozkłady empiryczne względnych błędów prognozy średniej stopy zwrotu oraz współczynnika Sharpe'a portfela wykazuje silną prawostronną asymetrię. Można w nich również odnaleźć kilka skrajnie dodatnio odstających obserwacji, co prawdopodobnie powoduje, iż odchylenie standardowe jest względnie wysokie. Histogramy częstości uzyskanych rozkładów zaprezentowano na rysunkach 2 i 3.



Rys. 2. Histogram częstości rozkładu empirycznego względnych błędów prognozy średniej stopy zwrotu portfela



Rys. 3. Histogram częstości rozkładu empirycznego względnych błędów prognozy współczynnika Sharpe'a portfela

W związku z faktem, iż wspomniane rozkłady są prawostronnie asymetryczne i asymetria jest znaczna, estymacja przedziału ufności średniego błędu prognozy przy pomocy wzoru (6) nie pozwala na uzyskanie wiarygodnego wyniku. Podjęto zatem próbę ustalenia prawdopodobieństwa, że wartości mierników efektywności uzyskane na zbiorach testowych są większe od wartości odpowiadających mierników otrzymanych na zbiorach uczących. Dla średniej stopy zwrotu z portfela prawdopodobieństwo to wyniosło nieco ponad 60%, natomiast dla współczynnika Sharpe'a nieznacznie powyżej 55%. Wyniki te nie pozwalają na jednoznaczne stwierdzenie, że zaprezentowana metoda portfelowa daje możliwość uzyskania zadowalających rezultatów inwestycyjnych. Jednakże w zestawieniu z faktem, iż ponad 95% względnych błędów prognozy ma wartość niższą niż 1, można przypuszczać, że zastosowana metoda optymalizacji portfela daje umiarkowanie satysfakcjonujące wyniki.

Podsumowanie

W artykule przedstawiona została możliwość adaptacji metody powtarzanej K-składowej walidacji krzyżowej w aspekcie badania efektywności inwestycyjnej metody optymalizacji portfela zaproponowanej przez Markowitza [1952]. Rezultaty przeprowadzonego badania empirycznego wydają się potwierdzać postawioną hipotezę, która mówi, że zastosowana procedura wtórnego próbkowania daje możliwość wnioskowania statystycznego bez stawiania założeń dotyczących postaci analitycznej rozkładu badanej cechy. Co więcej, metoda ta pozwala na uzyskanie rozkładów empirycznych o bardzo dużej liczebności nawet gdy niedostępna jest duża liczba obserwacji historycznych. W wypadku wnioskowania w odniesieniu do skuteczności metod inwestycyjnych jest to bardzo istotne, gdyż w wielu przypadkach uzyskanie jednej wiarygodnej obserwacji miernika efektywności wymaga co najmniej kilkuset obserwacji stóp zwrotu.

Literatura

- Bunke O., Droge B. (1984), *Bootstrap and Cross-Validation Estimates of the Prediction Error for Linear Regression Models*, "The Annals of Statistics", Vol. 12, No. 4, s. 1400-1424.
- Efron B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, "Society of Industrial and Applied Mathematics CBMS-NSF Monographs", Vol. 38.

- Fushiki T. (2011), *Estimation of Prediction Error by Using K-Fold Cross Validation*, "Statistics and Computing", Vol. 21, s. 137-146.
- Liu L., Ozsu T., red. (2009), *Encyclopedia of Database Systems*, Springer-Verlag, New York, s. 532-538.
- Markowitz H. (1952), *Portfolio Selection*, "The Journal of Finance" Vol. 7, No. 1.
- Sharpe W.F. (1994), *The Sharpe Ratio*, "The Journal of Portfolio Management", New York.
- Sobczyk M. (2000). *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa.
- Tibshirani R.J., Tibshirani R. (2009), *A Bias Correction for the Minimum Error Rate in Cross-Validation*, "Annals of Applied Statistics", Vol. 3, No. 2, s 822-829.

POTENTIAL APPLICATIONS OF RESAMPLING METHODS FOR PORTFOLIO INVESTMENTS EFFICIENCY VERIFICATION

Summary: This paper presents one of possible adaptations of resampling method – repeated K-fold cross-validation – used for portfolio optimization strategies investment efficiency estimation. This method is very helpful with statistical inference determination in situation when analytical form of distribution is not known or analytical form makes using available parametric methods inefficient. Empirical study was conducted to confirm that aforementioned assumption is true and its results show that there is no support for assertion of its falseness. It was performed on large historical data set from American stock market. Moreover, this article presents how to determine prediction errors for investment efficiency measures applied in resampling methods.

Keywords: K-fold cross validation, resampling, portfolio methods, investment efficiency, bootstrap.