

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 328

Taksonomia 23

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowci

Krzysztof Jajuga, Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2014

Redaktor Wydawnictwa: Barbara Majewska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

w Dolnośląskiej Bibliotece Cyfrowej www.dbc.wroc.pl,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się
na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2014

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	11
Małgorzata Rószkiewicz , Wykorzystanie metaanalizy w budowaniu modelu pomiarowego w przypadku braku niezmienniczości zasad pomiaru na przykładzie pomiaru zadowolenia z życia.....	13
Elżbieta Sobczak , Harmonijność inteligentnego rozwoju regionów Unii Europejskiej	21
Ewa Roszkowska, Renata Karwowska , Analiza porównawcza województw Polski ze względu na poziom zrównoważonego rozwoju w roku 2010.....	30
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Analiza porównawcza wybranych filtrów w analizie synchronizacji cyklu koniunkturalnego.....	41
Marcin Salamaga , Próba konstrukcji tablic „wymierania scenicznego” spektakli operowych na przykładzie Metropolitan Opera.....	51
Iwona Foryś , Wykorzystanie analizy dyskryminacyjnej do typowania rynków podobnych w procesie wyceny nieruchomości niemieszkalnych	59
Jerzy Korzeniewski , Selekcja zmiennych w klasyfikacji – propozycja algorytmu	69
Sabina Denkowska , Testowanie wielokrotne przy weryfikacji wieloczynnikowych modeli proporcjonalnego hazardu Coxa.....	76
Ewa Chodakowska , Teoria równań strukturalnych w klasyfikacji zmiennych jawnych i ukrytych według charakteru ich wzajemnych oddziaływań	85
Iwona Konarzewska , Model PCA dla rynku akcji – studium przypadku	94
Katarzyna Wójcik, Janusz Tuchowski , Dobór optymalnego zestawu słów istotnych w opiniach konsumentów na potrzeby ich automatycznej analizy	106
Aleksandra Łuczak , Zastosowanie metody AHP-LP do oceny ważności determinant rozwoju społeczno-gospodarczego w jednostkach administracyjnych	116
Aleksandra Witkowska, Marek Witkowski , Klasyfikacja pozycyjna banków spółdzielczych według stanu ich kondycji finansowej w ujęciu dynamicznym	126
Adam Depta , Zastosowanie analizy korespondencji do oceny jakości życia ludności na podstawie kwestionariusza SF-36v2	135
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Indukcja reguł dla danych niekompletnych i niezbalansowanych: modele klasyfikatorów i próba ich zastosowania do predykcji ryzyka operacyjnego w torakochirurgii	146

Małgorzata Misztal , Wybrane metody oceny jakości klasyfikatorów – przegląd i przykłady zastosowań.....	156
Anna M. Olszewska , Wykorzystanie wybranych metod taksonomicznych do oceny potencjału innowacyjnego województw	167
Iwona Bąk , Porównanie jakości grupowań powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej.....	177
Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Segmentacja gospodarstw domowych według wydatków na turystykę zorganizowaną.....	186
Agnieszka Wałęga , Podejście syntetyczne w analizie spójności ekonomicznej gospodarstw domowych.....	196
Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Zastosowanie analizy korespondencji do badania wpływu elektrowni wiatrowych na jakość życia ludności	205
Joanna Banaś, Krzysztof Małecki , Klasyfikacja punktów pomiarów ankietowych kierowców na granicy Szczecina z wykorzystaniem zmiennych symbolicznych.....	214
Aneta Becker , Wykorzystanie informacji granularnej w analizie wymagań rynku pracy.....	222
Katarzyna Cheba, Joanna Holub-Iwan , Wykorzystanie analizy korespondencji w segmentacji rynku usług medycznych.....	230
Adam Depta, Iwona Staniec , Identyfikacja czynników decydujących o jakości życia studentów łódzkich uczelni.....	238
Katarzyna Dębowska, Jarosław Kilon , Reguły asocjacyjne w analizie wyników badań metodą Delphi.....	247
Anna Domagała , O wykorzystaniu analizy głównych składowych w metodzie <i>Data Envelopment Analysis</i>	254
Alicja Grześkowiak , Analiza wykluczenia cyfrowego w Polsce w ujęciu indywidualnym i regionalnym.....	264
Anna M. Olszewska, Anna Gryko-Nikitin , Pomiar postrzegania jakości kształcenia uczelni wyższej na danych porządkowych z wykorzystaniem środowiska R.....	273
Karolina Paradysz , Hierarchiczna metoda grupowania powiatów jako podejście benchmarkowe w ocenie bezrobocia według BAEL-u w wybranych typach małych obszarów	282
Radosław Pietrzyk , Porównanie metod pomiaru efektywności zarządzania portfelami funduszy inwestycyjnych.....	290
Agnieszka Przedborska, Małgorzata Misztal , Wybrane metody statystyki wielowymiarowej w ocenie skuteczności terapeutycznej głębokiej stymulacji elektromagnetycznej u pacjentów z chorobą zwyrodnieniową stawów.....	299

Wojciech Roszka, Marcin Szymkowiak , Podejście kalibracyjne w statystycznej integracji danych	308
Iwona Skrodzka , Zastosowanie wybranych metod klasyfikacji do analizy kapitału ludzkiego krajów Unii Europejskiej	316
Agnieszka Stanimir , Wielowymiarowa analiza czynników sprzyjających włączeniu społecznemu	326
Dorota Strózik, Tomasz Strózik , Przestrzenne zróżnicowanie poziomu życia w województwie wielkopolskim.....	334
Izabela Szamrej-Baran , Identyfikacja przyczyn ubóstwa energetycznego w Polsce przy wykorzystaniu modelowania miękkiego.....	343
Janusz Tuchowski, Katarzyna Wójcik , Klasyfikacja obiektów w systemie Krajowych Ram Kwalifikacji opisanych za pomocą ontologii	353
Aleksandra Matuszewska-Janica , Grupowanie krajów Unii Europejskiej ze względu na poziom feminizacji sektorów gospodarczych	361
Monika Rozkrut, Dominik Rozkrut , Identyfikacja strategii innowacyjnych przedsiębiorstw usługowych w Polsce	369

Summaries

Małgorzata Rószkiewicz , The use of meta-analysis in building the measurement model in case of the absence of measurement invariance on the example of measuring of life satisfaction.....	20
Elżbieta Sobczak , Harmonious smart growth of European Union regions.....	29
Ewa Roszkowska, Renata Karwowska , The comparative analysis of Polish voivodeships with respect to sustainable development in 2010.....	40
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Comparative analysis of chosen filters in business cycles analysis	50
Marcin Salamaga , The attempt of construction of the life tables for opera works on the example of the Metropolitan Opera	58
Iwona Foryś , Using discriminant analysis to select similar markets in non-residential property valuation process.....	68
Jerzy Korzeniewski , Variable selection in classification – algorithm proposal	75
Sabina Denkowska , Multiple testing in the verification process of multifactorial Cox proportional hazards models	84
Ewa Chodakowska , The theory of structural equations modelling in the classification of observed variables and latent constructs according to the character of their relationship.....	93
Iwona Konarzewska , Modelling stock market by PCA factor model – case study	105

Katarzyna Wójcik, Janusz Tuchowski , Selection of the optimal set of relevant words in consumers opinions in the context of the opinion mining ..	115
Aleksandra Łuczak , Application of AHP-LP to the evaluation of importance of determinants of socio-economic development in the administrative units	125
Aleksandra Witkowska, Marek Witkowski , A dynamic approach to the ranking of cooperative banks by their financial condition	134
Adam Depta , Application of correspondence analysis for the measurement of quality of life – questionnaire SF-36v2 based research	145
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Classification rules extraction for missing and imbalance data: models of classifiers and initial results in the rules-based thoracic surgery risk prediction.....	155
Małgorzata Misztal , Selected methods for assessing the performance of classifiers – an overview and examples of applications.....	166
Anna M. Olszewska , The application of selected quantitative methods to the evaluation of voivodeship innovation level potential.....	176
Iwona Bąk , The comparison of the quality of groupings of poviats of West Pomeranian Voivodeship in terms of tourism attractiveness	185
Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Household segmentation with respect to the expenditure on organized tourism.....	195
Agnieszka Wałęga , Synthetic approach in the analysis of economic coherence of households	204
Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Using the correspondence analysis to examine the impact of wind turbines on the quality of life.....	213
Joanna Banaś, Krzysztof Małecki , Classification of measurement survey points of drivers on the boundary of Szczecin using symbolic variables...	221
Aneta Becker , The use granular information in the analysis of the requirements of the labor market.....	229
Katarzyna Cheba, Joanna Hołub-Iwan , The application of the correspondence analysis of patients segmentation on the medical service market	237
Adam Depta, Iwona Staniec , Identification of the factors that determine the quality of students life at universities in Lodz.....	246
Katarzyna Dębkowska, Jarosław Kilon , Association rules in the analysis of research results the Delphi method	253
Anna Domagała , About using Principal Component Analysis in Data Envelopment Analysis	263
Alicja Grześkowiak , Analysis of the digital divide in Poland at the individual and regional level	272

Anna M. Olszewska, Anna Gryko-Nikitin , Assessment of perception of quality of teaching at an institution of higher learning based on the ordinal data with the utilization of R environment.....	281
Karolina Paradysz , The hierarchical method of grouping poviats as a benchmark approach in the assessment of unemployment by BAEL in selected types of small areas	289
Radosław Pietrzyk , Comparison of methods of measuring the performance of investment funds portfolios.....	298
Agnieszka Przedborska, Małgorzata Misztal , Selected multivariate statistical analysis methods in the evaluation of efficacy of deep electromagnetic stimulation in patients with degenerative joint disease	307
Wojciech Roszka, Marcin Szymkowiak , A calibration approach in statistical data integration	315
Iwona Skrodzka , Application of some methods of classification to the analysis of human capital in the European Union.....	325
Agnieszka Stanimir , Multivariate analysis of social inclusion factors.....	333
Dorota Strózik, Tomasz Strózik , Spatial differentiation of the standard of living in Great Poland Voivodeship	342
Izabela Szamrej-Baran , Identification of fuel poverty causes in Poland using soft modelling	352
Janusz Tuchowski, Katarzyna Wójcik , Classification of objects in the National Classification Framework described by the ontology.....	360
Aleksandra Matuszewska-Janica , Clustering of European Union states taking into consideration the levels of feminization of economic sectors..	368
Monika Rozkrut, Dominik Rozkrut , Identification of service sector innovation strategies in Poland.....	379

Sabina Denkowska

Uniwersytet Ekonomiczny w Krakowie

TESTOWANIE WIELOKROTNE PRZY WERYFIKACJI WIELOCZYNNIKOWYCH MODELI PROPORCJONALNEGO HAZARDU COXA

Streszczenie: Jednym z etapów weryfikacji modelu proporcjonalnego hazardu Coxa jest badanie istotności statystycznej parametrów modelu, które zazwyczaj przeprowadza się, każde na poziomie istotności α , lekceważąc fakt testowania wielokrotnego, a to skutkuje błędnymi decyzjami o istotności w rzeczywistości nieistotnych czynników. W artykule zaproponowano wykorzystanie procedur testowań wielokrotnych do kontroli, czy pozostawione w modelu czynniki nie zostały uznane za istotne z powodu testowania wielokrotnego podczas weryfikacji ich istotności statystycznej.

Słowa kluczowe: model proporcjonalnego hazardu Coxa, testowanie wielokrotne, FWER, FDR.

1. Wstęp

Wielowymiarowe metody analizy przeżycia wypracowane na gruncie medycyny coraz powszechniej wykorzystywane są w badaniach ekonomicznych, socjologicznych czy technicznych (inżynieryjnych). Jedną z najpopularniejszych metod analizy przeżycia jest model proporcjonalnego hazardu Coxa. Jego popularność wynika z szeregu zalet, do których zaliczyć należy brak założeń dotyczących postaci rozkładu „przeżyć”. Zatem w przeciwieństwie do klasycznych modeli regresji, można go stosować w przypadku, gdy zmienna objaśniana nie ma rozkładu normalnego. Od klasycznych modeli regresji model Coxa odróżnia również możliwość uwzględnienia w nim danych niepełnych. Oszacowany model proporcjonalnego hazardu Coxa wykorzystywany jest do wyznaczenia współczynników ryzyka dla poszczególnych czynników, jak również do estymacji funkcji przeżycia. Niezwykle istotne jest zatem, by oszacowany model dobrze opisywał badaną zależność, a zawarte w nim czynniki nie były uznane za istotne statystycznie tylko z powodu braku kontroli efektu testowania wielokrotnego.

Jednym z etapów weryfikacji modelu proporcjonalnego hazardu Coxa jest badanie istotności statystycznej czynników hazardu, które przeprowadza się zarówno

za pomocą testu globalnego, jak i wnioskuje się o każdym czynniku oddzielnie. Poszczególne testowania zazwyczaj przeprowadza się każde na poziomie istotności α , ignorując fakt testowania wielokrotnego, a to skutkuje błędnymi decyzjami o istotności w rzeczywistości nieistotnych czynników. Jeśli rozważymy teoretycznie testowanie m prawdziwych, niezależnych hipotez zerowych, każdą na poziomie istotności α , to prawdopodobieństwo odrzucenia przynajmniej jednej prawdziwej hipotezy zerowej wynosi $1 - (1 - \alpha)^m$. W przypadku 10 niezależnych, prawdziwych hipotez zerowych testowanych na poziomie istotności 0,05 prawdopodobieństwo odrzucenia co najmniej jednej prawdziwej hipotezy wynosi aż 0,4. W praktyce bardzo rzadko mamy do czynienia z niezależnymi testowaniami, co utrudnia kontrolę efektu testowania wielokrotnego. Zastosowanie procedur testowań wielokrotnych pozwala skontrolować, czy pozostawione w modelu czynniki nie zostały uznane za istotne tylko z powodu efektu testowania wielokrotnego.

2. Model proporcjonalnego hazardu Coxa

Funkcja hazardu w modelu proporcjonalnego hazardu Coxa [1972] jest dana następującym wzorem:

$$h(t; x_1, x_2, \dots, x_m) = h_0(t) \exp(\beta_1 x_1 + \dots + \beta_m x_m),$$

gdzie: $h(t; x_1, x_2, \dots, x_m)$ – wynikowy hazard przy danych m zmiennych niezależnych x_1, x_2, \dots, x_m i odpowiednim czasie przeżycia,

$h_0(t)$ – hazard bazowy (odniesienia) lub zerowa linia hazardu,

$\beta_1, \beta_2, \dots, \beta_m$ – współczynniki modelu,

t – czas obserwacji.

Współczynniki $\beta_1, \beta_2, \dots, \beta_m$ estymowane są metodą *częściowej największej wiarygodności* (*partial maximum likelihood method*) zaproponowaną przez Coxa [1972].

Jednym z etapów weryfikacji oszacowanego modelu proporcjonalnego hazardu Coxa jest badanie istotności statystycznej współczynników regresji uwzględnionych w modelu. Etap ten polega na testowaniu m hipotez postaci:

$$H_{0,i}: \beta_i = 0 \text{ vs. } H_{A,i}: \beta_i \neq 0 \quad (i = 1, 2, \dots, m).$$

Weryfikację istotności parametru β_i przeprowadzić można za pomocą testu *częściowego ilorazu wiarygodności* (*partial likelihood ratio test*), testu Walda, *score testu*. Testy te generalnie dają zbliżone wyniki i prowadzą do tych samych wniosków na temat istotności statystycznej parametrów modelu [Hosmer, Lemeshow 1999, s. 108]. W pakietach *STATISTICA* i *R* istotność statystyczna parametrów w modelu proporcjonalnego hazardu Coxa weryfikowana jest za pomocą testu Walda.

3. Miary błędu I rodzaju dla zbioru wnioskowań

W literaturze na temat testowań wielokrotnych najczęściej wymianiane są dwie miary błędu I rodzaju dla zbioru wnioskowań. Miary te zdefiniowane są następująco:

- FWER (*Family Wise Error Rate*): $\text{FWER} = P(V \geq 1)$,
- FDR (*False Discovery Rate*):
$$\text{FDR} = \begin{cases} E\left(\frac{V}{R}\right) & \text{gdy } R > 0 \\ 0 & \text{gdy } R = 0 \end{cases},$$

gdzie V oznacza zmienną losową określającą liczbę prawdziwych hipotez zerowych odrzuconych w procesie testowania m hipotez zerowych, a R – zmienną losową określającą liczbę odrzuconych hipotez zerowych.

Kontrola FWER oznacza kontrolę prawdopodobieństwa odrzucenia co najmniej jednej prawdziwej hipotezy zerowej na z góry ustalonym poziomie α . W przypadku licznych zbiorów wnioskowań znacznie maleje moc procedur kontrolujących FWER i wówczas alternatywą jest kontrola FDR, czyli wartości oczekiwanej frakcji błędnych odrzuceń wśród wszystkich odrzuceń hipotez zerowych. W tym przypadku α oznacza akceptowany odsetek błędnych odrzuceń wśród wszystkich odrzuceń hipotez zerowych.

4. Procedury testowań wielokrotnych

Model proporcjonalnego hazardu Coxa jest specyficznym modelem regresji, gdyż w przeciwieństwie do klasycznych modeli regresji można w nim uwzględnić dane niepełne oraz stosować go w przypadku, gdy zmienna objaśniana nie ma rozkładu normalnego. Ta specyfika modelu Coxa poważnie ogranicza możliwości wykorzystania klasycznych metod wnioskowań wielokrotnych (takich jak przedziały ufności¹ Scheffégo [Domański, Pruska 2000, s. 200-201] do kontroli efektu testowania wielokrotnego podczas badania istotności statystycznej parametrów modelu. W takiej sytuacji kontrolę wspomnianych miar błędu I rodzaju w przypadku rodziny wnioskowań dotyczących istotności parametrów modelu proporcjonalnego hazardu Coxa mogą zapewnić uniwersalne brzegowe procedury testowań wielokrotnych [Denkowska 2011] oraz procedury łączne oparte na resamplingu², uwzględniające łączny rozkład statystyk testowych, dzięki czemu charakteryzują się większą mocą od procedur brzegowych.

¹ Wielokrotne przedziały ufności pozwalają wnioskować o liniowych współczynnikach regresji i ich liniowych zależnościach na łącznym, przyjętym z góry, poziomie ufności dla całej rodziny wnioskowań $1 - \alpha$.

² Zastosowanie resamplingu w testowaniu wielokrotnym umożliwia przeprowadzanie testowania wielokrotnego mimo braku znajomości struktury kowariancyjnej danych czy też braku normalności.

Wygodnym pojęciem w teorii testowań wielokrotnych są skorygowane prawdopodobieństwa testowe [patrz np.: Wright 1992; Denkowska 2012]. Mając wyznaczone skorygowane prawdopodobieństwa testowe, decyzję o ewentualnym odrzuceniu hipotezy zerowej podejmujemy porównując odpowiadające jej skorygowane prawdopodobieństwo testowe z przyjętym łącznym poziomem istotności dla całej rodziny wnioskowań.

4.1. Uniwersalne brzegowe procedury testowań wielokrotnych

Najstarszą, a zarazem najprostszą procedurą testowań wielokrotnych, gwarantującą kontrolę FWER, jest bardzo konserwatywna (małej mocy) procedura Bonferroniego, dla której skorygowane prawdopodobieństwa testowe wyznaczone są ze wzoru:

$$\tilde{p}_j = \min(mp_j; 1) \text{ dla } j = 1, \dots, m. \quad (1)$$

Konserwatyzm ten jest tym poważniejszy, im silniejsze są zależności pomiędzy statystykami testowymi lub im liczniejsza jest rodzina wnioskowań. Zstępującą modyfikacją procedury Bonferroniego jest uniwersalna procedura Holma, dla której skorygowane prawdopodobieństwa testowe wyznaczone są następująco:

$$\begin{aligned} \tilde{p}_{(1)} &= \min(mp_{(1)}; 1); \\ \tilde{p}_{(j)} &= \min(\max(\tilde{p}_{(j-1)}; (m-j+1)p_{(j)}); 1) \text{ dla } j = 2, \dots, m. \end{aligned} \quad (2)$$

Kontrolę miary FDR warto rozważyć w przypadku dużych zbiorów złożonych z setek, tysięcy wnioskowań. Przeprowadzać ją można za pomocą uniwersalnej procedury Benjaminiego-Yekutieli [Benjamini, Yekutieli 2001], dla której skorygowane prawdopodobieństwa testowe wyznaczamy ze wzorów:

$$\begin{aligned} \tilde{p}_{(m)} &= \min\left(1; p_{(m)} \sum_{i=1}^m \frac{1}{i}\right); \\ \tilde{p}_{(m-j)} &= \min\left(\tilde{p}_{(m-j+1)}, p_{(m-j)} \frac{m}{m-j} \sum_{i=1}^m \frac{1}{i}\right) \text{ dla } j = 1, \dots, m-1. \end{aligned} \quad (3)$$

Zaletą brzegowych procedur testowań wielokrotnych jest ich prostota obliczeniowa.

4.2. Procedury łączne testowań wielokrotnych oparte na repróbkowaniu

W 1993 r. Westfall i Young [1993] zaproponowali procedury łączne testowań wielokrotnych oparte na maksimach statystyk testowych lub minimach prawdopodobieństw testowych. Procedury te zapewniają kontrolę FWER. Poważną wadą tych procedur jest wymóg „obrotowości podzbioru” (*subset pivotality*) [Westfall, Young 1993; Denkowska 2012], który niestety nie jest spełniony w sytuacji testowania

istotności współczynników regresji [Dudoit, van der Laan 2008]. Nie można ich zatem zastosować do kontroli testowania wielokrotnego przy badaniu istotności współczynników regresji w modelu proporcjonalnego hazardu Coxa.

Ciekawą alternatywę dla procedur Westfalla i Younga stanowią procedury, wolne od ograniczającego wymogu obrotowości podzbioru, zaproponowane przez Dudoit i van der Laana [2008]. Procedury Dudoit i van der Laana są zaimplementowane w pakiecie *multtest* w R pod nazwą MTP. Według dokumentacji pakietu *multtest* zapewniają kontrolę wybranej miary błędu I rodzaju dla rodziny wnioskowań (FWER, gFWER [patrz np. Denkowska 2012], FDR) w wielu sytuacjach badawczych, np. przy porównywaniu parami wartości przeciętnych, testowaniu istotności współczynników korelacji czy też testowaniu istotności współczynników regresji w modelu proporcjonalnego hazardu Coxa.

5. Przykład empiryczny

Sposoby kontroli efektu testowania wielokrotnego zaprezentowane zostaną na przykładzie przedstawionym u Hosmera i Lemeshowa [1999]. Badanie³ dotyczyło osób uzależnionych od narkotyków i miało na celu porównanie skuteczności dwóch terapii o różnym czasie trwania, prowadzonych jednocześnie w dwóch ośrodkach. Zmienna TIME określała czas (w dniach), jaki upłynął od zakończenia (przerwania) terapii do powrotu do nałogu. W badaniu uwzględniono dane niepełne.

W finalnym modelu proporcjonalnego hazardu Coxa zaprezentowanym u Hosmera i Lemeshowa [1999, s. 230, tab. 6.6] znalazły się następujące zmienne:

AGE – wiek (w latach),

BECKTOTA – liczba punktów w skali depresji Becka (od 0,00 do 54,00),

$NDRUGFP1 = [(NDRUGTX + 1)/10]^{-1}$,

$NDRUGFP2 = [(NDRUGTX + 1)/10]^{-1} \cdot \ln[(NDRUGTX + 1)/10]$,

gdzie NDRUGTX – liczba wcześniejszych terapii (0, ..., 40),

IVHX_3 – przyjmowanie narkotyków dożylnie (1 – ostatnio, 0 – wcześniej lub nigdy)⁴,

RACE – rasa (0 – biały, 1 – pozostałe),

TREAT – randomizowany wybór terapii (0 – krótka, 1 – długa),

SITE – miejsce leczenia (0 – ośrodek A, 1 – ośrodek B).

Ostateczną wersję tego modelu proporcjonalnego hazardu Coxa przedstawia tab. 1. Istotność statystyczną parametrów modelu badano na poziomie $\alpha = 0,1$.

³ W badaniu wykorzystano dane udostępnione na stronie: ftp://ftp.wiley.com/public/sci_tech_med/survival/.

⁴ Wyznaczona na podstawie zmiennej IVHX – historia przyjmowania narkotyków dożylnie (0 – nigdy, 1 – wcześniej, 3 – ostatnio).

Tabela 1. Finalna wersja modelu proporcjonalnego hazardu Coxa

Zmienne	Beta	Błąd standardowy	Statystyka Walda	p_i
AGE	-0,04139	0,009913	17,43733	0,0000298
BECKTOTA	0,00874	0,004965	3,09640	0,0784754
NDRUGFP1	-0,57442	0,125189	21,05407	0,0000045
NDRUGFP2	-0,21457	0,048587	19,50193	0,0000101
IVHX_3	0,22774	0,108563	4,40047	0,0359369
RACE	-0,46685	0,134754	12,00254	0,0005320
TREAT	-0,24674	0,094340	6,84069	0,0089146
SITE	-1,31691	0,531440	6,14048	0,0132173
AGExSITE	0,03240	0,016081	4,05916	0,0439406
RACExSITE	0,85022	0,247758	11,77624	0,0006007

Źródło: opracowane na podstawie Hosmer, Lemeshow [1999, s. 230, tab. 6.6] oraz obliczeń własnych w *STATISTICA* w celu poprawy precyzji wyników (na podstawie danych udostępnionych na ftp://ftp.wiley.com/public/sci_tech_med/survival/).

Do kontroli efektu testowania wielokrotnego przy weryfikacji istotności statystycznej parametrów zastosowano uniwersalne, brzegowe procedury testowań wielokrotnych: procedurę Bonferroniego, procedurę Holma oraz procedurę Benjamini-Yekutieli. Otrzymane wyniki przedstawiono w tab. 2.

Tabela 2. Uporządkowane prawdopodobieństwa testowe oraz odpowiadające im skorygowane prawdopodobieństwa testowe wyznaczone przez zastosowanie procedur Bonferroniego, Holma oraz Benjamini-Yekutieli

Zmienne	$p(i)$	Proc. Bonferroniego	Proc. Holma	Proc. B-Y
NDRUGFP1	0,0000045	0,000045	0,000045	0,000132
NDRUGFP2	0,0000101	0,000101	0,000091	0,000148
AGE	0,0000298	0,000298	0,000238	0,000291
RACE	0,0005320	0,005320	0,003724	0,003519
RACExSITE	0,0006007	0,006007	0,003724	0,003519
TREAT	0,0089146	0,089146	0,044573	0,043518
SITE	0,0132173	0,132173	0,052869	0,055304
IVHX_3	0,0359369	0,359369	0,107811	0,131573
AGExSITE	0,0439406	0,439406	0,107811	0,143001
BECKTOTA	0,0784754	0,784754	0,107811	0,229852

Źródło: obliczenia własne (w tabeli pogrubiono skorygowane prawdopodobieństwa testowe większe od 0,1).

Analiza skorygowanych prawdopodobieństw testowych pokazała, że bez względu na to, czy zdecydujemy się na kontrolę FWER, czy na bardziej liberalne podejście polegające na kontroli FDR, to nie można uznać wszystkich parametrów

modelu za istotne statystycznie. Zatem po usunięciu zmiennej z największym prawdopodobieństwem testowym należy kontynuować proces budowy modelu proporcjonalnego hazardu Coxa.

W tab. 3 przedstawiono model otrzymany po usunięciu zmiennej BECKTOTA.

Tabela 3. Wyniki estymacji parametrów modelu Coxa, po usunięciu zmiennej BECKTOTA

Zmienne	Beta	Błąd standardowy	Statystyka Walda	p_i
AGE	-0,04229	0,00976	18,7891	0,0000083
NDRUGT1	-0,55011	0,12337	19,8815	0,0000210
NDRUGT2	-0,20383	0,04790	18,1054	0,0200154
IVHX_3	0,24671	0,10606	5,4111	0,0005027
RACE	-0,45908	0,13193	12,1080	0,0084959
TREAT	-0,24452	0,09291	6,9267	0,0144030
SITE	-1,28585	0,52544	5,9887	0,0514656
AGExSITE	0,03107	0,01595	3,7933	0,0003641
RACExSITE	0,86909	0,24377	12,7109	0,0000083

Źródło: obliczenia własne.

Oszacowany model ponownie poddano analizie pod kątem kontroli efektu testowania wielokrotnego. W tab. 4 przedstawiono uzyskane wyniki.

Tabela 4. Uporządkowane prawdopodobieństwa testowe (dla modelu bez zmiennej BECKTOTA) oraz odpowiadające im skorygowane prawdopodobieństwa testowe wyznaczone przez zastosowanie procedur Bonferroniego, Holma oraz Benjaminiego-Yekutieli

Zmienne	$p_{(i)}$	Proc. Bonferroniego	Proc. Holma	Proc. B-Y
NDRUGFP1	0,0000083	0,000075	0,000074	0,000013
NDRUGFP2	0,0000146	0,000131	0,000117	0,000013
AGE	0,0000210	0,000189	0,000147	0,000013
RACE	0,0003641	0,003277	0,002184	0,000129
RACExSITE	0,0005027	0,004524	0,002514	0,000129
TREAT	0,0084959	0,076463	0,033983	0,001593
SITE	0,0144030	0,129627	0,043209	0,002058
IVHX_3	0,0200154	0,180139	0,043209	0,063701
AGExSITE	0,0514656	0,463190	0,051466	0,145594

Źródło: obliczenia własne (w tabeli pogrubiono skorygowane prawdopodobieństwa testowe większe od 0,1).

Analiza wartości skorygowanych prawdopodobieństw testowych otrzymanych dla procedury Bonferroniego w tabelach 2 oraz 4 wskazuje na poważny konserwatyzm tej metody. W tab. 4 procedura Bonferroniego wskazała aż trzy parametry jako nieistotne statystycznie, a procedura Holma, również gwarantująca kontrolę

FWER, pozwoliła uznać wszystkie parametry modelu za istotne statystycznie na poziomie 0,1. Co więcej, również procedura Benjaminiego-Yekutieliiego, kontrolująca FDR, wskazała jeden parametr jako nieistotny statystycznie i tym samym wypadła gorzej od procedury Holma. Kontrolę FDR za pomocą procedury Benjaminiego-Yekutieliiego warto jednak rozważyć w przypadku bardzo licznych zbiorów wnioskowań, gdy weryfikowana jest istotność setek czy nawet tysięcy parametrów modelu, bowiem wówczas wyraźnie maleje moc procedur kontrolujących FWER.

Kontrola efektu testowania wielokrotnego za pomocą procedur łącznych Dudoit i van der Laana [2008] niestety nie powiodła się. Funkcja MTP (z parametrem $test="coxph.Yvs.XZ"$) w wersji 2.16.0 pakietu *multtest* w R nie działa poprawnie dla modelu Coxa z kilkoma zmiennymi objaśniającymi.

6. Podsumowanie

W zaprezentowanym przykładzie zastosowanie prostych obliczeniowo, brzegowych procedur testowań wielokrotnych pozwoliło stwierdzić, że zmienna BECKTOTA niesłusznie została uznana za istotną statystycznie, a tym samym budowa modelu powinna być kontynuowana. Rezultat ten otrzymano na podstawie analizy finalnego modelu proporcjonalnego hazardu Coxa, bez ingerowania w proces budowy modelu. Fakt ten pokazuje konieczność kontroli efektu testowania wielokrotnego istotności parametrów modelu przynajmniej na końcowym etapie weryfikacji modelu. Zastosowanie prostych brzegowych procedur testowań wielokrotnych pozwala skontrolować, czy pozostawione w modelu parametry nie zostały uznane za istotne tylko z powodu efektu testowania wielokrotnego podczas weryfikacji ich istotności statystycznej. Należy mieć na uwadze, że ignorowanie faktu testowania wielokrotnego ma poważne konsekwencje – prowadzi do podejmowania błędnych decyzji o istotności, w rzeczywistości nieistotnych parametrów w modelu proporcjonalnego hazardu Coxa. Niestety, nie powiodła się próba wykorzystania łącznych procedur testowań wielokrotnych do kontroli efektu testowania wielokrotnego w modelu proporcjonalnego Coxa. Badanie pokazało, że stosunkowo nowe, „obiecujące” procedury łączne Dudoit i van der Laana [2008] w wersji zaimplementowanej w R pod nazwą MTP (w pakiecie *multtest* ver.2.16.0), zastosowane do testowania istotności parametrów w modelu proporcjonalnego hazardu Coxa z kilkoma zmiennymi objaśniającymi nie działają poprawnie.

Literatura

- Benjamini Y., Hochberg Y. (1995), *Controlling the false discovery rate: A practical and powerful approach to multiple testing*, „Journal of the Royal Statistical Society”, Ser. B, vol. 57, no.1, s. 289-300.
- Benjamini Y., Yekutieli D. (2001), *The Control of the False Discovery Rate in Multiple Testing Under Dependency*, „Annals of Statistics”, 29.

- Cox D.R. (1972), *Regression Models and Life-Tables*, „Journal of the Royal Statistical Society”, Series B (Methodological), Vol. 34, No. 2.
- Denkowska S. (2011), *Testowanie jednoczesne przy weryfikacji ocen parametrów strukturalnych modelu ekonometrycznego*, Zeszyty Naukowe UEK, Metody Analizy Danych, nr 873, Kraków, s. 53-68.
- Denkowska S. (2012), *Procedury testowań wielokrotnych*, [w:] K. Jajuga, M. Walesiak (red.), *Klasyfikacja i analiza danych – teoria i zastosowania*, Taksonomia 20, Prace Naukowe UE we Wrocławiu nr 278, s. 362-369.
- Domański C., Pruska K. (2000), *Nieklasyczne metody statystyczne*, PWE, Warszawa.
- Dudoit S., van der Laan M. (2008), *Multiple Testing Procedures with Applications to Genomics*, Springer Series in Statistics.
- Hochberg Y., Tamhane A.C. (1987), *Multiple Comparison Procedures*, John Wiley & Sons, NY 1987.
- Hosmer D.W., Jr. Lemeshow S. (1999), *Applied Survival Analysis. Regression Modeling of Time to Event Data*, John Wiley and Sons, New York.
- Westfall P.H., Young S.S. (1993), *Resampling Based Multiple Testing*, Wiley, New York.
- Wright S.P. (1992), *Adjusted P-values for Simultaneous Inference*, Biometrics, 48.

MULTIPLE TESTING IN THE VERIFICATION PROCESS OF MULTIFACTORIAL COX PROPORTIONAL HAZARDS MODELS

Summary: One of the stages of verifying Cox proportional hazards model is testing statistical significance of its parameters, each at the level α , disregarding the fact of multiple testing. Disregarding it results in taking wrong decisions concerning statistical significance of factors, which are, in fact, insignificant. Applying the procedures of multiple testing allows for confirmation whether the remaining factors have been regarded as significant only because of multiple testing during the verification of their statistical significance.

Keywords: Cox proportional hazards model, multiple testing, FWER, FDR.