



JAK ROZUMIEĆ I WYKONYWAĆ ANALIZĘ PRZEŻYCIA

Andrzej Sokołowski, Centrum Onkologii – Oddział w Krakowie, Uniwersytet Ekonomiczny w Krakowie, StatSoft Polska

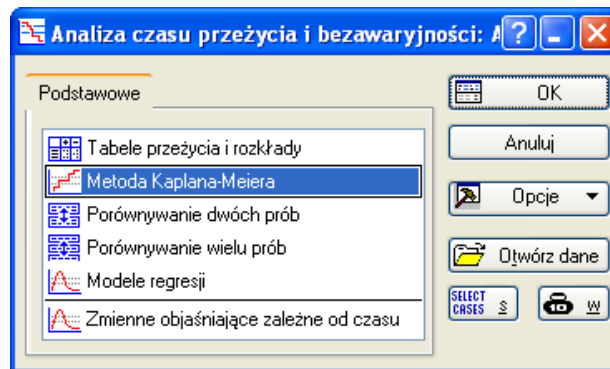
Analiza czasu przeżycia to inaczej analiza czasu trwania. W procesie „trwania” w czasie jednostki statystycznej mamy zdefiniowane dwa zdarzenia, przy czym pierwsze z nich musi nastąpić wcześniej niż drugie. Okres czasu, jaki upływa pomiędzy tymi zdarzeniami jest zmienną losową. W życiu każdego człowieka takimi zdarzeniami są narodziny i śmierć. Jest oczywiste, że długość życia zależy od wielu czynników, w tym również od czynników losowych. Inne przykłady zagadnień, w których analizujemy czas trwania zjawiska, to: czas bezawaryjnej pracy urządzenia, czas pozostawania na bezrobociu, czas istnienia przedsiębiorstwa itp. W przedstawianych tu zastosowaniach analizy przeżyć zakładamy, że każde z dwóch zdarzeń wyznaczających czas trwania może w danej jednostce statystycznej zajść tylko jeden raz. Ten warunek jest w oczywisty sposób spełniony przy analizie czasu trwania życia organizmu żywego. Tutaj tym drugim zdarzeniem, tzw. *punktem końcowym* jest śmierć. Jeżeli inaczej zdefiniujemy taki punkt końcowy, na przykład jako *przerzut nowotworu*, to zdarzenie takie może występować więcej niż jeden raz, w różnych momentach. W takich przypadkach za moment kończący *przeżycie* przyjmujemy pierwsze wystąpienia zdarzenia. Najbardziej charakterystyczną cechą typowych danych, jakimi posługujemy się w analizie przeżyć, jest obecność obiektów, w których końcowe zdarzenie nastąpiło (wówczas mamy do czynienia z obserwacją *kompletną*), oraz obiektów, w których to zdarzenie (jeszcze) nie nastąpiło (obserwacja *ucięta*). Ta specyficzność danych statystycznych doprowadziła do powstania specjalnych metod stosowanych tylko w analizie czasu trwania zjawisk.

Szacowanie funkcji przeżyć

Jak wspomnieliśmy, cechą charakterystyczną danych do analizy przeżyć jest występowanie obserwacji uciętych. Terminu *censoring* (*ucinięcie, obcinanie, cenzurowanie*) użył po raz pierwszy Hald w 1949 roku.

W trakcie przygotowania danych należy zadbać o precyzyjne zdefiniowanie tzw. „punktu końcowego”, czyli zdarzenia, które kończy okres przeżycia. Przy przeżyciach całkowitych (*OS – overall survival*) takim zdarzeniem jest zgon. Podobnym przeżyciem jest przeżycie do zgonu wywołanego badaną chorobą (*DSS – disease specific survival*). Zgon jest wówczas zaliczany jako zdarzenie kończące przeżycie, jeżeli jego przyczyną była analizowana

choroba. Zgon z innych przyczyn powoduje, że mamy do czynienia z obserwacją uciętą. Dane o przeżyciach muszą zawierać dwie kolumny. Pierwsza z nich (właściwe przeżycie) podaje czas przeżycia (najczęściej wyrażany w miesiącach). Druga zmienna (zwana zmienną wskaźnikową) zawiera informację o stanie pacjenta „na końcu” okresu przeżycia. Zmienna ta jest kodowana dwiema wartościami. W zasadzie mogą to być dowolne liczby naturalne. Jednak wydaje się, że rozsądną konwencją jest oznaczanie 1 sytuacji, gdy zdarzenie kończące przeżycie nastąpiło (jest to obserwacja kompletna, gdyż przeżycie już nie może być dłuższe), zaś przez 0 sytuacji, gdy to zdarzenie nie nastąpiło (jest to obserwacja ucięta, gdyż potencjalnie „przeżycie” biegnie nadal). Jeżeli pacjent został stracony z obserwacji, to jego przeżycie jest liczone do momentu ostatniej informacji o nim, i jeżeli zdarzenie nie nastąpiło, to w drugiej zmiennej (wskaźnikowej) wpisujemy zero.



Rys. 1. Okno analizy przeżycia.

Pierwsza opcja dostępna w oknie *Analizy czasu przeżycia i bezawaryjności* to tzw. *tablice trwania życia*. Technika oparta na tablicach trwania życia jest jedną z najstarszych metod analizy danych dotyczących przeżyć. Tablicę taką można traktować jako rozbudowaną tablicę rozkładu liczebności. Rozkład czasów przeżycia dzieli się na pewną liczbę przedziałów. Dla każdego przedziału możemy obliczyć liczbę i proporcję przypadków lub obiektów, które weszły do danego przedziału „żywe”, liczbę i proporcję przypadków, które uległy awarii w danym przedziale (tzn. liczbę ostatecznych zdarzeń lub liczbę przypadków, które „wymarły”) oraz liczbę przypadków utraconych lub uciętych w danym przedziale. Metoda ta jest powszechnie stosowana w demografii, gdzie przedziały przyjmowane są naturalnie jako okresy roczne lub pięcioletnie.

W statystyce medycznej jest już praktycznie nieużywana. Zamiast klasyfikować obserwowane przeżycia w postaci tablicy trwania życia możemy estymować funkcję przeżycia bezpośrednio z ciągłych czasów przeżycia. Wyobraźmy sobie, że tworzymy tablicę trwania życia tak, że każdy przedział czasowy zawiera dokładnie jeden przypadek. Wymnażając prawdopodobieństwa przeżycia z tych „przedziałów” (tzn. dla każdej pojedynczej obserwacji), otrzymujemy funkcję przeżycia, pozwalającą ocenić prawdopodobieństwo przeżycia określonej liczby miesięcy, dni czy lat (w zależności od tego, jakiej jednostki czasu używamy w badaniach). Estymator Kaplana-Meiera wykorzystuje oczywiste spostrzeżenie, że aby przeżyć na przykład 12 miesięcy, należy przeżyć najpierw pierwszy miesiąc, następnie drugi, potem, trzeci, czwarty itd., i wreszcie dwunasty. Poszukiwaną

ocenę prawdopodobieństwa przeżycia dwunastu miesięcy znajdujemy, mnożąc kolejne prawdopodobieństwa warunkowe według wzoru:

$$P(\text{przeżycia 12 miesięcy}) = P(\text{przeżycia pierwszego miesiąca}) * P(\text{przeżycia drugiego miesiąca pod warunkiem, że przeżyło się pierwszy miesiąc}) * P(\text{przeżycia trzeciego miesiąca pod warunkiem, że przeżyło się dwa pierwsze}) * \dots * P(\text{przeżycia dwunastego miesiąca pod warunkiem, że przeżyło się jedenaście poprzednich})$$

Prawdopodobieństwa składające się na powyższy iloczyn szacuje się osobno i to umożliwia wykorzystanie obserwacji uciętych.

Przewaga metody Kaplana-Meiera nad metodą tablic trwania życia polega na tym, że oceny funkcji przeżycia nie zależą od grupowania danych (na pewną liczbę przedziałów czasowych). Metoda Kaplana-Meiera i metoda tablic trwania życia dają identyczne oceny, jeśli przedziały tablicy trwania życia zawierają najwyżej po jednej obserwacji.

Rozważania ilustrowane będą danymi dotyczącymi kobiet chorych na raka piersi. Wszystkie z nich były operowane. Większość poddana została chemioterapii, radioterapii i hormonoterapii. Mamy dane o 172 pacjentkach. Jak zwykle w plikach medycznych, dane te są niekompletne, co nie przeszkadza w poprawnym prowadzeniu analiz. Zazwyczaj mamy tu trzy możliwości: usunięcie pacjentki z brakami danych ze wszystkich analiz, usunięcie pacjentki z tych analiz, które dotyczą zmiennych brakiem danych, oraz zastępowanie brakujących danych średnią. Najbardziej godne polecenia jest wyjście drugie, a najmniej – trzecie. W omawianym przykładzie analizować będziemy przeżycie bezobjawowe.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	IMIE	Wiek	Liczba węzłów	T	Bloom	Typ operacji	Początek RT	Koniec RT	Dawka	Przeżycia całkowite	Endpoint PC	Przeżycie bezobjawowe	Endpoint PB	P53 (%)	P53(%) kod	wiek <=50	Liczba węzłów - klasy	Czas RT	RT > 30 dni
1	Teresa	32		2	3	Patey	2002/04/02	2002/04/08	5000	52,0	0	52,0	0	0,4	0	1		6	0
2	Alina	54		2	3	Patey				27,5	0	15,0	1	21,2	1	0			
3	Anna	41		2	2	Patey								0,0	0	1			
4	Ewa	57		2		Patey								0,6	0	0			
5	Grażyna	48		1		Patey	2003/02/25	2003/04/01	5000					0,0	0	1		35	1
6	Józefa	55		2	2	Patey	2001/03/21	2001/04/25	5000							1	0	35	1
7	Grażyna	46		2	3	Patey				35,0	0	28,0	1	80,3	1	1			
8	Danuta	58		2		Patey	2001/08/28	2001/10/01	5000	18,0	0	15,0	1	0,1	0	0		34	1
9	Ilona	35	5	2	3	Patey	2001/03/23	2001/04/30	5000	47,0	0	21,0	1	1,6	1	1		38	1

Rys. 2. Nagłówek bazy danych.

Analiza przeżycia znajduje się w programie *STATISTICA* w grupie metod *Zawansowane modele liniowe i nieliniowe*. Po wybraniu *Metody Kaplana-Meiera* trzeba wskazać zmienne, które zawierają czas przeżycia oraz informację o tym, czy dana obserwacja jest kompletna czy ucięta. Popularna jest konwencja oznaczania obserwacji kompletnej przez 1 („zdarzenie nastąpiło”), a obserwacji uciętej przez 0 („zdarzenie nie nastąpiło”).

Omawiany estymator funkcji przeżycia po raz pierwszy został zaproponowany przez E.L. Kaplana i P. Meiera w 1958 roku (choć pierwszą ideę tego pomysłu można znaleźć w pracy P.E. Böhmera z 1912 roku). Estymator Kaplana-Meiera jest wyrażony wzorem:

$$\hat{S}(t) = \prod_{t_j \leq t} (1 - \frac{d_j}{r_j}),$$



gdzie Π jest symbolem iloczynu, r_j – liczbą narażonych w okresie t_j , a d_j – liczbą „zdarzeń” (zgonów) w tym okresie.

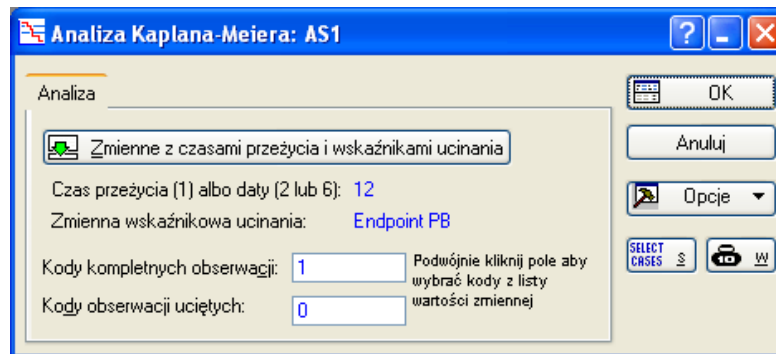
Wariancja tego estymatora jest określona przez wzór Greenwooda (1926):

$$\hat{\sigma}^2(t) = [\hat{S}(t)]^2 \sum_{t_j \leq t} \frac{d_j}{r_j(r_j - d_j)}$$

Dla dużych prób rozkład estymatora Kaplana-Meiera zmierza do rozkładu normalnego, co pozwala na wyznaczenie przedziału ufności, według wzoru:

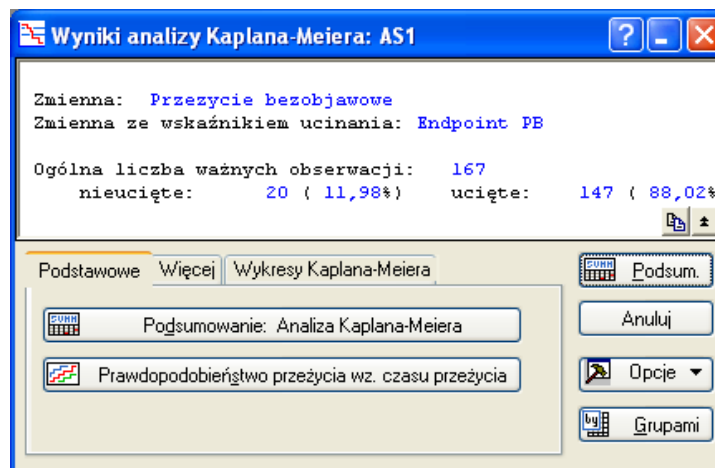
$$\hat{S}(t) \pm u_{1-\alpha/2} \hat{\sigma}(t).$$

Estymator Kaplana-Meiera inaczej nazywany jest granicznym estymatorem iloczynowym (*Product Limit Estimator*, w skrócie PL), ponieważ szukana ocena prawdopodobieństwa przeżycia jest iloczynem prawdopodobieństw warunkowych przeżycia kolejnych okresów czasu.



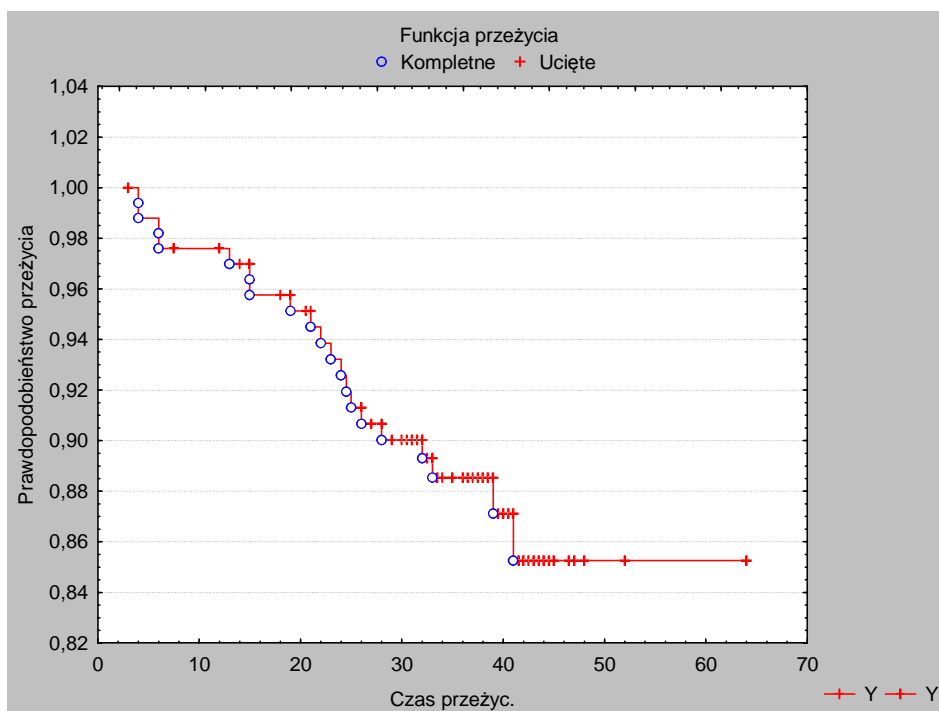
Rys. 3. Okno metody Kaplana-Meiera.

Najważniejsze wyniki zawarte są na karcie *Podstawowe*. Wcześniej warto zwrócić uwagę na łączną liczbę kompletnych obserwacji. Nie powinna ona być zbyt mała.



Rys. 4. Okno wyników szacowania funkcji przeżycia.

Pod klawiszem *Podsumowanie: Analiza Kaplana-Meiera* otrzymujemy tabelę ze współzrędnymi opisującymi krzywą przeżyć. Z tabeli tej można ocenić przeżycia w dowolnym momencie okresu obserwacji. Drugi klawisz wywołuje jeden z podstawowych rysunków wykorzystywanych analizie przeżyć. Jest to krzywa przeżyć, która podaje prawdopodobieństwo zgonu (lub innego zdarzenia) w poszczególnych jednostkach czasu. Powszechnie używane określenie „krzywa” jest tu niezbyt precyzyjne. Krzywą jest funkcja przeżyć dla populacji, natomiast ocena uzyskana metodą Kaplana-Meiera jest funkcją schodkową. „Schodki” występują w momentach czasu, w których nastąpiło przynajmniej jedno zdarzenie. Funkcja nie zmienia swego przebiegu, jeżeli jakiś pacjent jest w danym momencie czasu „stracony z obserwacji”. Rysunek surowy, jaki proponuje program, wymaga zazwyczaj dalszej obróbki dla uzyskania ilustracji możliwej do zaakceptowania przez wydawnictwo.

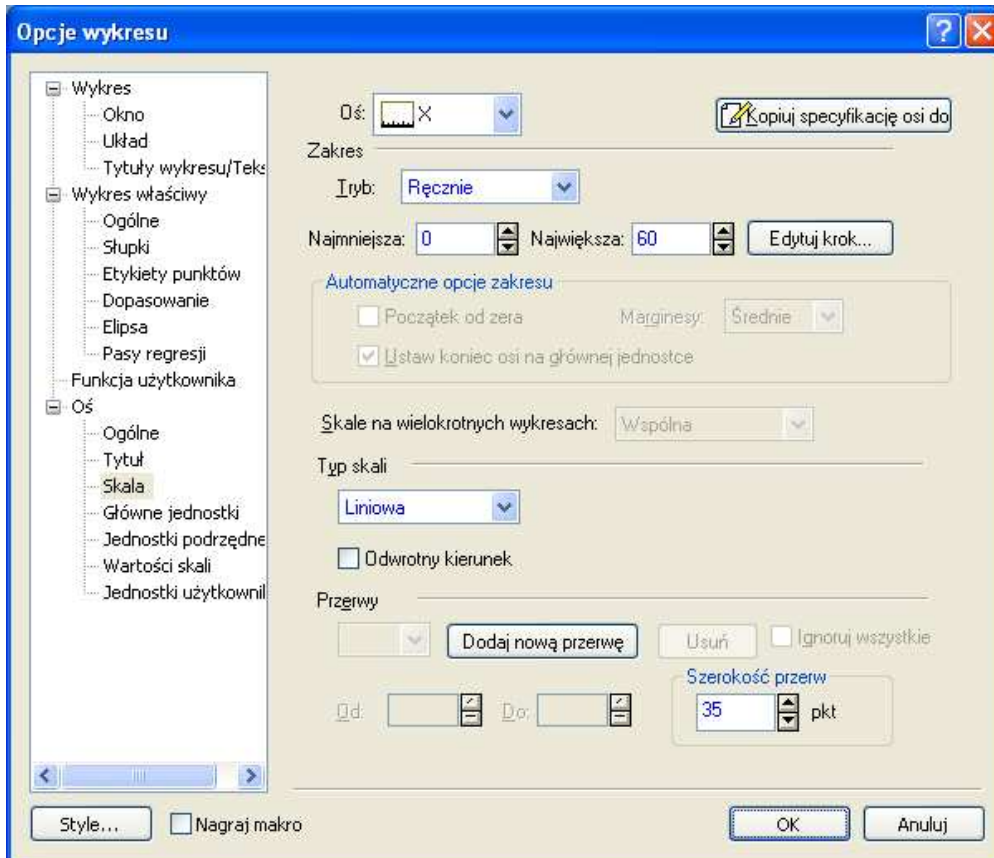


Rys. 5. Surowy rysunek funkcji przeżycia.

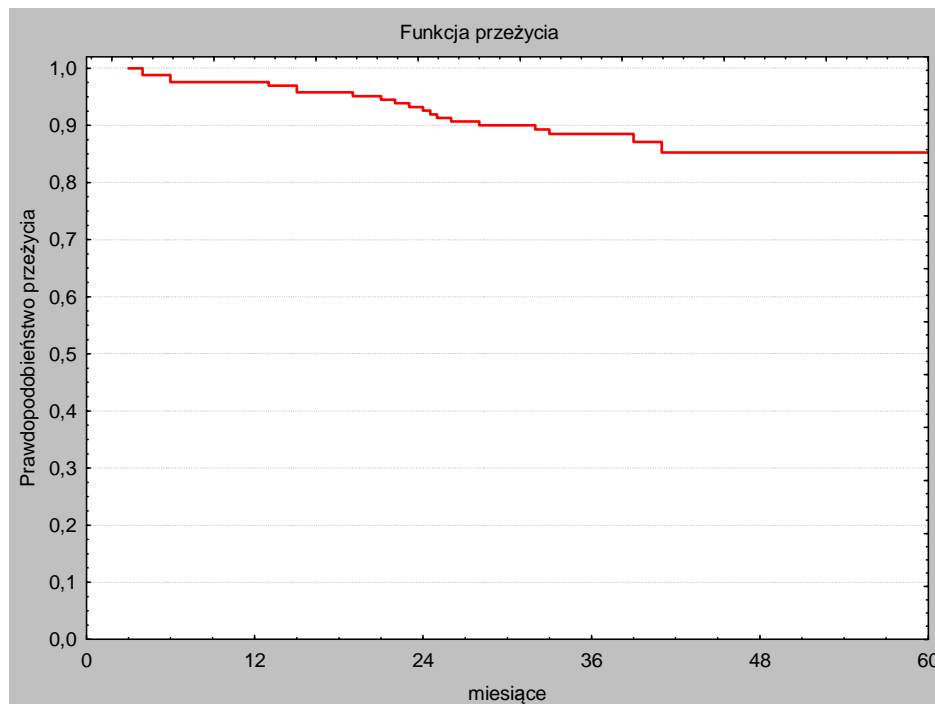
Klikając dwukrotnie w pole obramowania wykresu, uzyskujemy okno pozwalające zredagować wykres zgodnie z wymaganiami.

Akceptowalna rycina, przedstawiająca krzywą przeżyć bezobjawowych, pokazana jest jako rys. 7.

Pod klawiszem *Percentyle funkcji przeżycia* można uzyskać wartości mediany przeżyć oraz wartości dolnego i górnego kwartyła przeżyć. Te wszystkie wartości można też odczytać bezpośrednio z tablicy podającej pełne współrzędne funkcji przeżycia. Trzeba pamiętać, że przy dobrej przeżywalności możemy mieć do czynienia z sytuacją, w której oszacowanie mediany przeżyć jest niemożliwe. Dzieje się tak wtedy, gdy w okresie objętym obserwacją zmarło mniej niż 50% pacjentów. Można wtedy stwierdzić jedynie, że „mediana przeżyć jest większa od ...”.



Rys. 6. Okno opcji wykresu.



Rys. 7. Krzywa przeżyć bezobjawowych dla 172 chorych na raka piersi.

Wartości funkcji przeżycia w konkretnych momentach czasu można odczytać z danych wykresu. W tym celu ustawiamy kursor na obramowaniu wykresu właściwego (kolor szary) i przyciskamy prawy klawisz myszy. Pojawia się rozwijalne menu, z którego wybieramy opcję *Edytor danych wykresu*. Otrzymujemy tabelę, która w kolumnie *Schodkowy* zawiera informacje o przeżyciach. Te same dane uzyskamy, naciskając klawisz *Podsumowanie: Analiza Kaplana-Meiera*.

Przyp. Liczba	Czas	Skumul. Przeżyc.	Błąd Stand.
126+	3,00000		
143	4,00000	0,993976	0,006006
59	4,00000	0,987952	0,008468
62	6,00000	0,981928	0,010339
30	6,00000	0,975904	0,011902
10+	7,50000		
39+	12,00000		
116	13,00000	0,969804	0,013299
12+	14,00000		
2	15,00000	0,963666	0,014563
8	15,00000	0,957528	0,015710
106+	15,00000		
78+	15,00000		
32+	18,00000		
11	19,00000	0,951270	0,016808
40+	19,00000		
99+	20,50000		

Rys. 8. Wartości funkcji przeżyć bezobjawowych dla 172 chorych na raka piersi.

Wartości funkcji przeżycia odczytujemy w kolumnie *Skumul. Przeżyc.* Jeżeli dla danego momentu czasowego nie ma informacji, to przeżycie odczytujemy z wartości powyżej. Przykładowo przeżycie 12-miesięczne wynosi w naszym przykładzie 0,976. Zazwyczaj podaje się je nie jako prawdopodobieństwo, tylko w procentach – tutaj jako 97,6%. Przeżycia dla kolejnych lat można przedstawić w następującej tabeli.

Lata	1	2	3	4	5
Przeżycie	97,6	92,6	88,5	85,3	85,3

Porównywanie przeżyć

Przeżycia można porównywać w dwóch lub więcej próbach. Zasadniczo - ponieważ czasy przeżycia nie podlegają rozkładowi normalnemu - powinno się stosować testy nieparametryczne, które są oparte na porządku rangowym czasów przeżycia. W module *Testy nieparametryczne i rozkłady* znajduje się obszerny zestaw testów nieparametrycznych,



które można wykorzystać w celu porównania czasów przeżycia; jednakże testy w tamtym module nie mogą „operować” obserwacjami uciętymi.

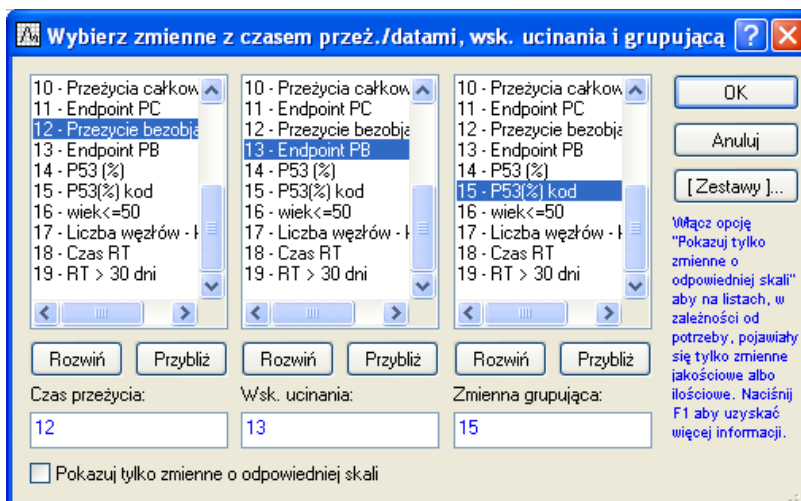
Analiza przeżycia zawiera pięć różnych (najczęściej nieparametrycznych) testów dla danych uciętych: Gehana uogólnienie testu Wilcozona, test Coxa-Mantela, test F Coxa, test log-rank oraz wersja Peto i Peto testu Wilcozona. Dostępny jest także test nieparametryczny dla porównania wielu grup. Większości tych testów towarzyszą odpowiednie wartości statystyk podlegających rozkładowi normalnemu standaryzowanemu. Wartości te można wykorzystać do testowania istotności statystycznej wszelkich różnic między grupami. Pamiętajmy, że większość tych testów daje rzetelne wyniki tylko przy odpowiednio dużych próbach; „zachowanie” się testów przy małych próbach nie zostało dokładnie zbadane.

Omówiona poprzednio metoda Kaplana-Meiera służy do szacowania pojedynczej krzywej przeżycia. W tej części prezentujemy testy wykorzystywane do porównywania krzywych oszacowanych metodą Kaplana-Meiera. Te krzywe przeżyć dotyczą podgrup pacjentów. Podgrupy te zdefiniowane są jakąś cechą jakościową. Do analizy potrzebna jest więc taka grupująca zmienna jakościowa. W programie *STATISTICA* mamy osobną część, która pozwala na porównanie dwóch grup (na przykład porównanie kobiet z mężczyznami, czy porównanie pacjentów, u których zadany parametr jest w normie, z tymi, którzy są poza normą), oraz następną, która umożliwia porównanie więcej niż dwóch grup (część ta może też być wykorzystywana do porównywania tylko dwóch grup).

Testy dla dwóch prób

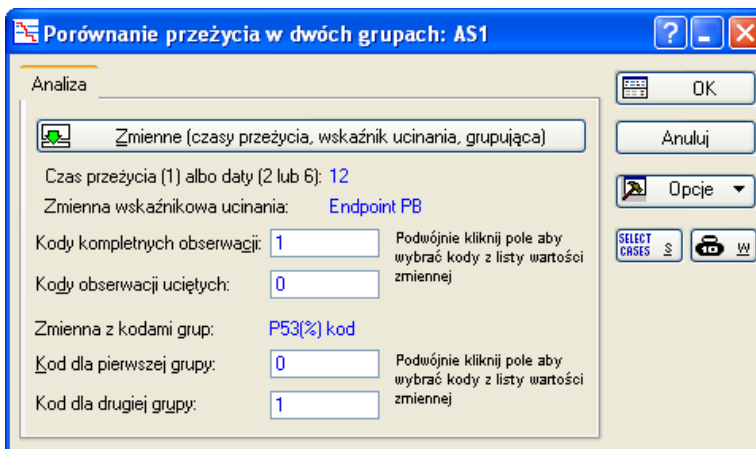
Po wybraniu *Porównania dwóch prób* pojawia się okno *Porównanie przeżycia w dwóch grupach* (program podaje tu też nazwę pliku), z którego należy przejść do określenia zmiennych potrzebnych do wykonania analizy.

W naszym przykładzie o raku piersi zbadamy wpływ p53 na przeżycie bezobjawowe. Białko p53 jest jednym z czynników transkrypcyjnych odgrywających ważną rolę podczas aktywacji zaprogramowanej śmierci komórki. Proces apoptozy pozwala na utrzymanie homeostazy w organizmie. Jednakże w komórkach rakowych p53 jest nieaktywne, co pozwala na dalszy rozwój nowotworu. Gen kodujący p53 jest zmutowany w ponad połowie wszystkich rodzajów nowotworów. Mutacja ta jest nieco inna niż w przypadku pozostałych białek zmutowanych w komórkach rakowych. Przeważnie zmiany w takich białkach są bardzo szerokie, np. delecja fragmentu genu lub mutacja typu „frame shift”. Białko takie jest niezdolne do pełnienia jakiegokolwiek funkcji. Natomiast w przypadku p53 jest to mutacja punktowa i polega na substytucji kilku par zasad. Wynikiem mutacji jest białko, które ma zmienione najważniejsze aminokwasy odpowiadające za wiązanie DNA. Tak więc mutacja ta jest dość „skromna”, a jednocześnie zupełnie wystarczająca, aby dezaktywować p53. Tetramery, w skład których wchodzi chociaż jeden zmutowany polipeptyd p53 są niezdolne do wiązania DNA [Vousden i Lu (2002)]. W naszym pliku danych mamy informację o procencie zmutowanego p53 oraz zmienną zerojedynkową oznaczającą te chore, dla których ten procent jest większy od 1,5.

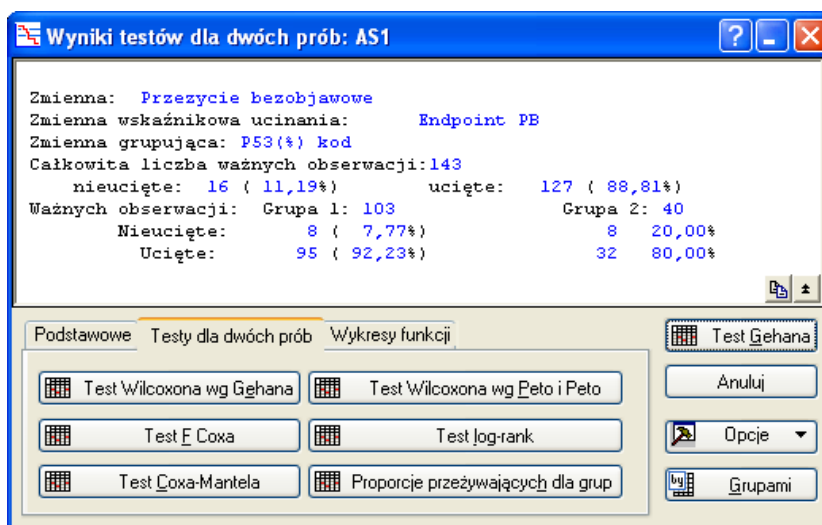


Rys. 9. Wybór zmiennych do porównania przeżyć w dwóch grupach.

Po wyborze zmiennych nasze „ustawienia” pojawiają się w oknie porównań.



Rys. 10. Okno zmiennych wybranych do porównania przeżyć.



Rys. 11. Okno zmiennych wybranych do porównania przeżyć.



Po naciśnięciu klawisza OK i przejściu na kartę *Testy dla dwóch prób* otrzymujemy okno widoczne na rys. 11.

Nie ma powszechnie uznanych wskazówek mówiących, którego testu należy użyć w danej sytuacji. Test F Coxa wydaje się być mocniejszy niż Gehana uogólnienie testu Wilcoxon, gdy:

1. próby są małe (tzn. na grupę mniejsze niż 50);
2. próby pochodzą z rozkładu wykładniczego lub rozkładu Weibulla;
3. nie ma obserwacji uciętych.

Lee, Desu i Gehan (1975) porównali test Gehana z kilkoma alternatywnymi procedurami i pokazali, że test Coxa-Mantela i test log-rank są mocniejsze (bez względu na obcinanie), gdy próby zostały pobrane z populacji o rozkładzie wykładniczym lub Weibulla. Jednocześnie przy takich założeniach występuje mała różnica między testem Coxa-Mantela a testem log-rank.

Najpopularniejszym testem jest niewątpliwie test log-rank. Przedstawimy go bliżej.

Czytając literaturę dotyczącą analizy przeżycia, można dojść do wniosku, iż opiera się ona na trzech głównych filarach. Są nimi: estymator Kaplana-Meiera, test log-rank oraz model proporcjonalnego hazardu Coxa.

Test log-rank inaczej, zwany też testem Mantela-Coxa, to nieparametryczny test statystyczny porównujący rozkład przeżycia (funkcję przeżycia) dwóch lub więcej niezależnych prób. Jak w przypadku estymatora Kaplana-Meiera znowu są w mocy założenia dotyczące braku informacji o cenzurowaniu (obserwacje mogą być prawostronnie cenzurowane).

Założmy na początek, że mamy dwie grupy danych o defektach (lub śmierci na skutek zdarzenia), które łączymy i budujemy dla nich przedziały podobne do tych dla estymatora Kaplana-Meiera. Jako d_{1i} i d_{2i} zapiszemy liczbę zdarzeń (zgonów) poszczególnych grup w przedziale $[0, t_1)$, $[t_1, t_2)$, ... itd. Podobnie jak poprzednio r_{1i} , r_{2i} to liczba przypadków w każdej z grup, która dożyła początku przedziału i , a w_{1i} , w_{2i} to liczba utraconych z obserwacji.

$$r_{1,i+1} = Y_{1,i} - d_{1,i} - w_{1,i}$$

$$r_{2,i+1} = Y_{2,i} - d_{2,i} - w_{2,i}$$

$$d_i = d_{1i} + d_{2i}, \quad r_i = r_{1i} + r_{2i}$$

Jeśli momenty obserwacji określonego zjawiska nie powtarzają się, wtedy dla pojedynczego przedziału $d_{1i} = 1, d_{2i} = 0 \vee d_{1i} = 0, d_{2i} = 1$. Dla każdego przedziału i możemy stworzyć tabelę przypadków 2x2, podobną do tej:

	Zdarzenie (zgon)	Brak zdarzenia	Sumy
Grupa 1	d_{1i}	$r_{1i} - d_{1i}$	r_{1i}
Grupa 2	d_{2i}	$r_{2i} - d_{2i}$	r_{2i}
Sumy	d_i	$r_i - d_i$	



Hipoteza zerowa dla testu zakłada brak różnic w rozkładzie przeżycia w obu grupach. Sprawdzenie jej wymaga spojrzenia na różnice w liczbie obserwacji w każdej grupie, które doświadczyły zdarzenia w poszczególnych przedziałach względem ich spodziewanej liczby. Przyjmujemy założenie o stałości sum marginalnych we wspomnianych tablicach przypadków. Pozostałe wartości poszczególnych komórek mogą zostać ustalone, gdy znamy jedynie d_{1i} , którą traktujemy jako zmienną losową. Hipoteza zerowa mówi, iż prawdopodobieństwo zdarzenia (śmierci) członka Grupy 1 w momencie t_i jest takie samo jak dla członka Grupy 2 i równe p . Stąd też prawdopodobieństwo obserwacji d_{1i} defektów w pierwszej grupie i d_{2i} defektów w drugiej, przy ogólnej liczby defektów równej d_i dla momentu t_i , wynosi:

$$P(d_{1i}|d_i r_{1i} r_{2i}) = \frac{\binom{r_{1i}}{d_{1i}} p^{d_{1i}} (1-p)^{r_{1i}-d_{1i}} \binom{r_{2i}}{d_{2i}} p^{d_{2i}} (1-p)^{r_{2i}-d_{2i}}}{\binom{r_i}{d_i} p^{d_i} (1-p)^{r_i-d_i}} = \frac{\binom{r_{1i}}{d_{1i}} \binom{r_{2i}}{d_{2i}}}{\binom{r_i}{d_i}}$$

Jak widać z powyższego wzoru, p daje się skrócić. Wyprowadzenie wzoru na średnią i wariancję d_{1i} dokonuje się za pomocą funkcji tworzących i założeniu hipergeometrycznego rozkładu tej zmiennej. Ostatecznie poszczególne wzory wyglądają następująco:

$$e_{1i} = \frac{r_{1i} d_i}{r_i}$$

$$v_{1i} = \frac{r_{1i} r_{2i} d_i (r_i - d_i)}{r_i^2 (r_i - 1)}$$

Statystyka testowa jest porównywana przy danym poziomie istotności z rozkładem chi kwadrat o jednym stopniu swobody (przypadek dwóch grup) i wylicza się ją według wzoru:

$$W = \frac{(\sum_{i=1} d_{1i} - \sum_{i=1} e_{1i})^2}{\sum_{i=1} v_{1i}}$$

Odrzucamy hipotezę zerową, jeśli $W > \chi_1^2$. Jeśli zechcielibyśmy wykonać test dla r -grup ($r > 2$) z hipotezą zerową mówiącą o braku różnic między funkcjami przeżycia w poszczególnych grupach, to statystyka testowa uproszczonego testu log-rank liczona będzie z wzoru:

$$W = \sum_{k=1}^r \frac{(\sum_{i=1} d_{ki} - \sum_{i=1} e_{ki})^2}{\sum_{i=1} e_{ki}}$$

gdzie

$$e_{ki} = \frac{r_{ki} d_i}{r_i}$$

Tym razem statystyka ma rozkład chi kwadrat o $r-1$ stopni swobody.

Jeszcze jedna forma tego testu, która uwzględnia podział grup na wewnętrzne warstwy. Warstwy te mogą wynikać z podziału obserwacji względem wieku, płci, czy też miejsca zamieszkania. Hipoteza zerowa zakłada brak różnicy między grupami pod względem rozkładu przeżycia dla każdej warstwy. Kroki postępowania w trakcie przygotowania do obliczenia statystyki testowej mogą przebiegać następująco:

- ◆ Etap 1 – podział danych na warstwy (*stratum*),
- ◆ Etap 2 – przygotowanie tablicy przypadków dla każdej z warstw $l = 1, \dots, s$,
- ◆ Etap 3 – obliczenie dla każdej z warstw:

$$O_{kl} = \sum_{i=1}^r d_{lki} - \text{zaobserwowana liczba zdarzeń w warstwie } l, \text{ grupy } k$$

$$E_{kl} = \sum_{i=1}^r e_{lki} - \text{spodziewana liczba zdarzeń w warstwie } l, \text{ grupy } k$$

$$e_{lki} = \frac{r_{lki}d_{li}}{r_{li}}$$

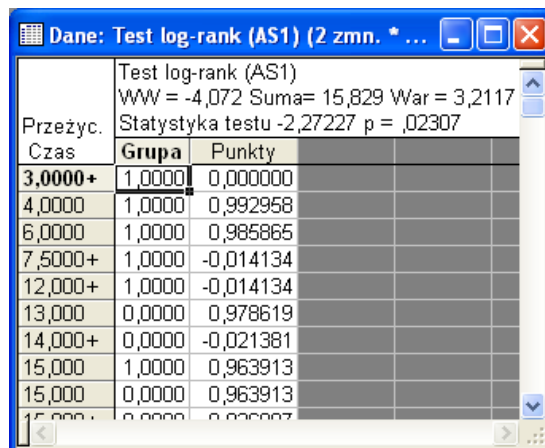
- ◆ Etap 4 – połączenie wyników z każdej tabeli w statystyce testowej:

$$W = \sum_{k=1}^r \frac{(\sum_{l=1}^s (O_{kl} - E_{kl}))^2}{\sum_{l=1}^s E_{kl}}$$

Podobnie jak dla testu dla wielu grup porównujemy tę statystykę z wartościami krytycznymi rozkładu χ^2_{r-1} .

Test log-rank może być wykorzystywany do znajdowania „optymalnej” kategoryzacji zmiennych ciągłych. Procedura ta spotyka się z krytyką wielu statystyków, którzy uważają, że są to zmienne sztucznie „istotne” i niesłusznie faworyzowane potem w konfrontacji ze zmiennymi „rzeczywistymi”.

Wracamy do naszego przykładu. Po wybraniu *Testu log-rank* otrzymujemy następujące okno wyników:

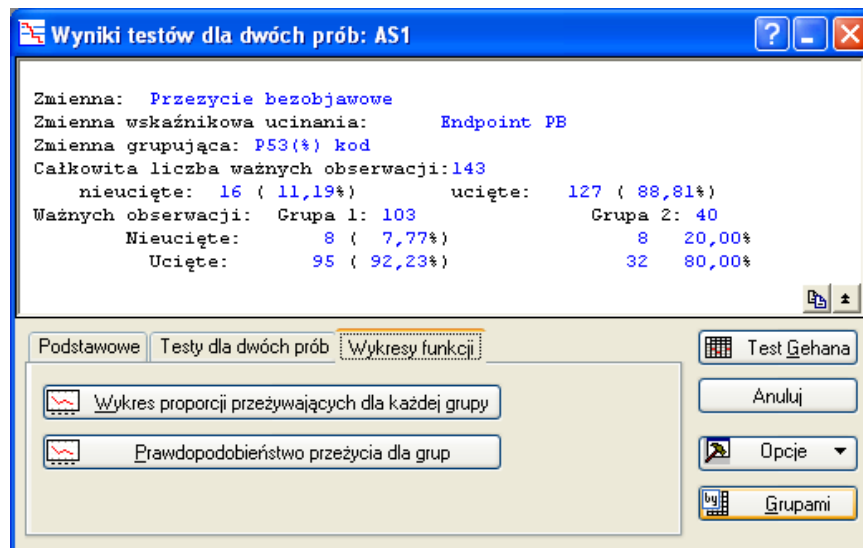


Przeżyc. Czas	Grupa	Punkty
3,0000+	1,0000	0,000000
4,0000	1,0000	0,992958
6,0000	1,0000	0,985865
7,5000+	1,0000	-0,014134
12,000+	1,0000	-0,014134
13,000	0,0000	0,978619
14,000+	0,0000	-0,021381
15,000	1,0000	0,963913
15,000	0,0000	0,963913
15,000	0,0000	0,963913

Rys. 12. Wyniki testu log-rank przy porównywaniu dwóch grup.



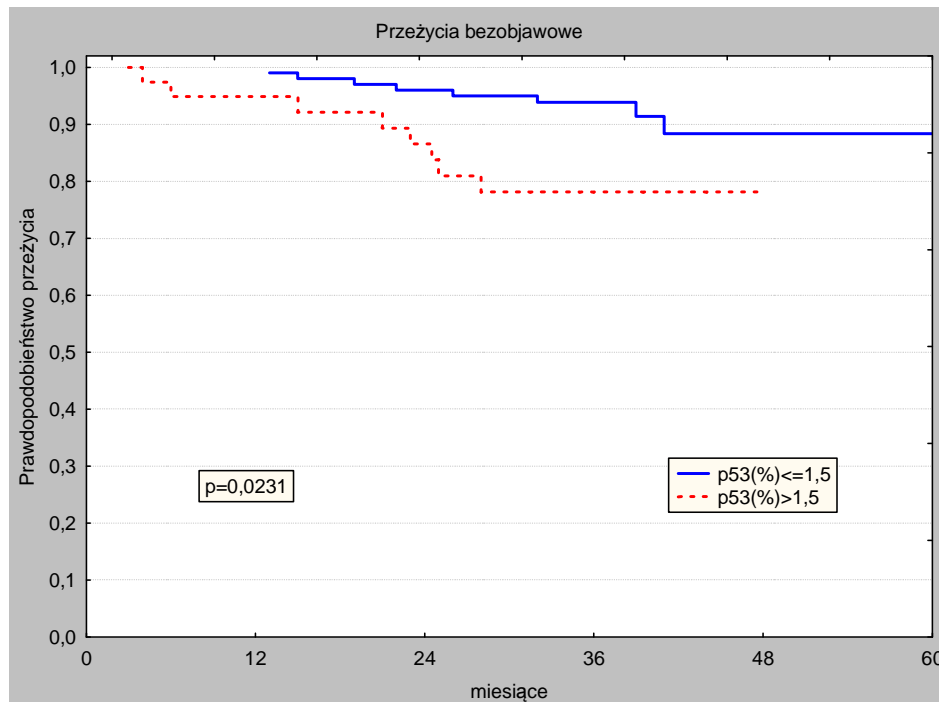
Najważniejszą informacją jest tu wartość p . Wynosi ona 0,0231, co świadczy o istotnej różnicy w przeżyciu bezobjawowym grup z $p53(\%)>1,5$ i $p53(\%)<=1,5$. Właściwą ilustracją jest tu rysunek, który uzyskujemy z karty *Wykresy funkcji*.



Rys. 13. Okno wyboru wykresu.

Wybieramy *Prawdopodobieństwo przeżycia dla grup*.

Wykres – po odpowiednich przeróbkach redakcyjnych – wygląda następująco:



Rys. 14. Przeżycia bezobjawowe w zależności od p53.

Korzystając z opcji *Edytor danych wykresu*, uzyskujemy dane potrzebne do przygotowania prostej tabelki podsumowującej przeżycia.

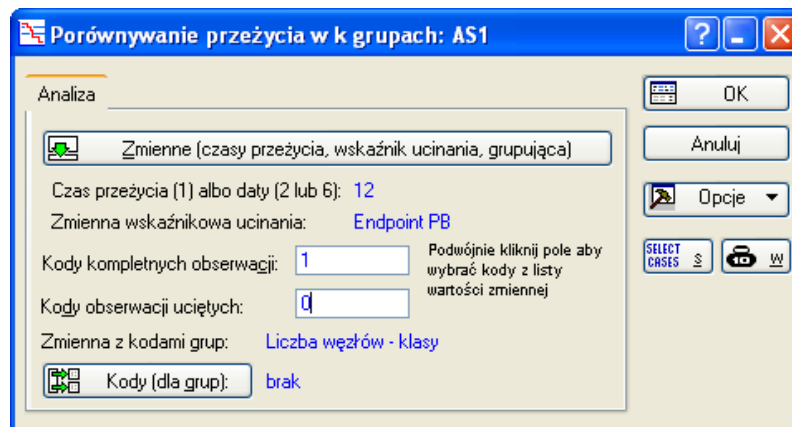
Lata	1	2	3	4	5
p53(%)<=1,5	100,0	96,0	93,9	88,4	88,4
p53(>1,5	94,9	86,6	78,1	78,1	.

Test dla wielu prób

Test dla wielu prób zastosowany w *Analizie przeżycia* jest rozwinięciem (lub uogólnieniem) testu Gehana, testu Peto i Peto oraz testu log-rank. Najpierw przy zastosowaniu procedury Mantela każdemu czasowi przeżycia przypisuje się punkty; następnie oblicza się wartość statystyki chi-kwadrat w oparciu o sumy (dla każdej grupy) tych punktów. Jeśli mamy tylko dwie grupy, to test ten jest równoważny uogólnieniu Gehana testu Wilcozona i w tym wypadku obliczany będzie domyślnie ten test.

Jak zwykle przy porównywaniu przeżyć, musimy określić zmienną zawierającą czas przeżycia, zmienną wskaźnikową informującą, czy obserwacja ma charakter kompletny czy ucięty, oraz zmienną, w której podana jest kategoryzacja chorych na grupy.

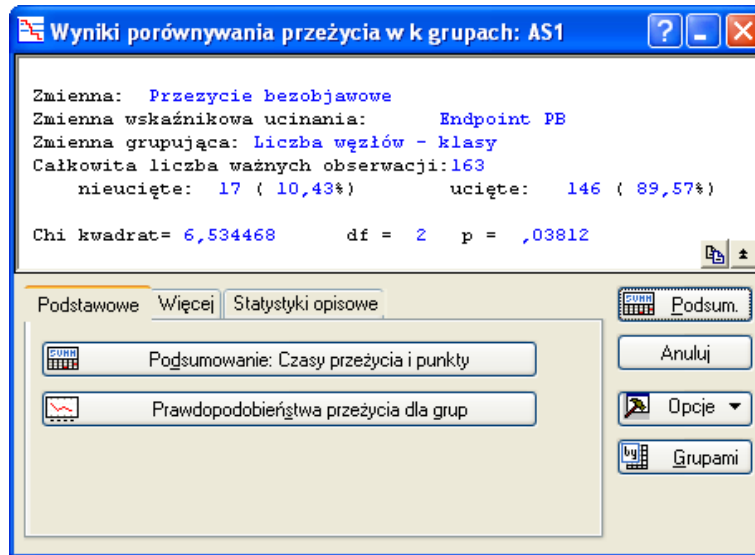
W przykładzie dotyczącym raka piersi rozważymy zależność przeżyć od liczby węzłów zmienionych przerzutowo. Pierwotnie ta liczba węzłów jest – jak wspomnieliśmy – zmienną skokową, mierzalną, podającą liczbę węzłów z meta. Dla celów tej analizy obszar zmienności tej cechy podzielimy na trzy klasy: klasa 1 – liczba zajętych węzłów do trzech, klasa 2 – liczba zajętych węzłów od czterech do 10, klasa 3 – ponad dziesięć zajętych węzłów.



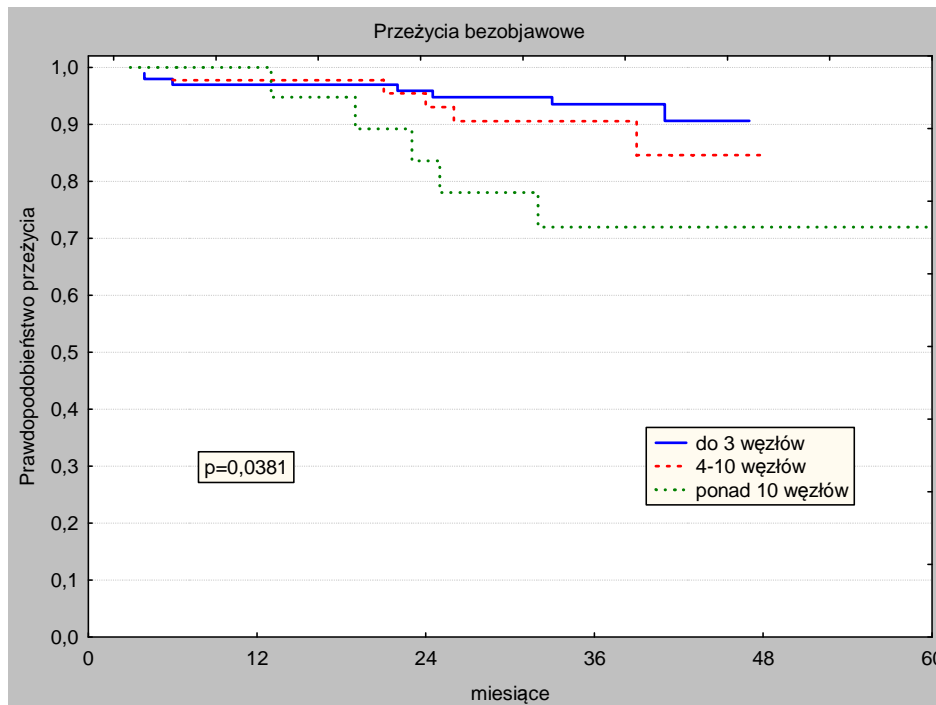
Rys. 15. Wybór zmiennych do porównania przeżyć w wielu grupach.

Wyniki testowania otrzymujemy po naciśnięciu OK.

Wartość $p=0,0381$ wskazuje na różnicę istotną statystycznie. Rysunek uzyskujemy z klawisza *Prawdopodobieństwa przeżycia dla grup*. Po obróbce redakcyjnej pokazuje on wyraźne różnice w przeżyciach trzech wyróżnionych grup.



Rys. 16. Wyniki porównania przeżyć bezobjawowych.



Rys. 17. Krzywe przeżyć w zależności od liczby zajętych węzłów chłonnych.

Przedstawione w artykule metody analizy przeżyć oraz ich wyniki miały przede wszystkim na celu zilustrowanie najpopularniejszych technik wykorzystywanych w badaniach medycznych i dostępnych w programie *STATISTICA 9*.



Literatura

1. Böhmer P.E. (1912), „Theorie der unabhängigen Wahrscheinlichkeiten”, *Rapports, Mémoires et Procèsverbaux de Septième Congrès International d’Actuaries*, Amsterdam, 2, 327-343.
2. Greenwood M. (1926), „The Natural Duration of Cancer”, *Reports of Public Health and Medical Subjects*, No. 33, London: H.M. Stationery Office.
3. Hald A. (1949). „Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point”, *Skandinavisk Aktuarietidskrift*, 1949, 119-134.
4. Kaplan E.L., Meier P. (1958), „Nonparametric estimation from incomplete observations”, *Journal of the American Statistical Association*, 53, 457-481.
5. Lee E.T., Desu M.M., Gehan E.A. (1975), „A Monte Carlo study of the power of some two-sample tests”, *Biometrika*, 62 (2), 425-532.
6. Vousden K.H., Lu X. (2002), „Live or let die: the cell’s response to p53”, *Nature*, vol. 2, 594-604.