

*IBM SPSS Complex Samples 29*



**Uwaga**

Przed użyciem tych informacji i produktu, którego one dotyczą, przeczytaj informacje znajdujące się w sekcji “Uwagi” na stronie 41.

**Informacje o produkcie**

Niniejsze wydanie dotyczy wersji 29, wydania 0, modyfikacji 1 produktu IBM® SPSS Statistics oraz wszystkich kolejnych wydań i modyfikacji, dopóki nie zostanie to określone inaczej w nowych wydaniach.

© Copyright International Business Machines Corporation .

---

# Spis treści

<b>Rozdział 1. Próby złożone.....</b>	<b>1</b>
Wprowadzenie do procedur prób złożonych.....	1
Własności prób złożonych.....	1
Użycie procedur prób złożonych.....	2
Dalsza lektura.....	2
Próbkowanie z projektu złożonego.....	2
Tworzenie nowego planu losowania.....	3
Kreator losowania prób: zmienne projektowe.....	3
Kreator losowania prób: metoda losowania.....	4
Kreator pobierania próbek: wielkość próby.....	4
Kreator losowania prób: zmienne wyników.....	5
Kreator losowania próbek: podsumowanie planu.....	6
Kreator próbkowania: opcje wyboru losowania próbek.....	6
Kreator losowania prób: pliki wyjściowe prób losowych.....	6
Kreator próbkowania: zakończenie.....	7
Modyfikowanie istniejącego planu losowania.....	7
Kreator losowania próbek: podsumowanie planu.....	7
Wykonywanie istniejącego planu losowania.....	7
Dodatkowe funkcje komend CSPLAN i CSSELECT.....	7
Przygotowywanie próbki złożonej do analizy.....	8
Tworzenie nowego planu analizy.....	8
Kreator przygotowania analizy: zmienne projektowe.....	8
Kreator przygotowania analizy: metoda szacowania.....	9
Kreator przygotowania analizy: wielkość.....	9
Kreator przygotowania analizy: podsumowanie planu.....	10
Kreator przygotowania analizy: zakończenie.....	10
Modyfikowanie istniejącego planu analizy.....	11
Kreator przygotowania analizy: podsumowanie planu.....	11
Plan prób złożonych.....	11
Częstości dla prób złożonych.....	11
Statystyki częstości dla prób złożonych.....	12
Brakujące wartości dla prób złożonych.....	12
Opcje dla prób złożonych.....	13
Próby złożone - Statystyki opisowe.....	13
Statystyki opisowe dla prób złożonych.....	14
Brakujące wartości w statystykach opisowych dla prób złożonych.....	14
Opcje dla prób złożonych.....	14
Tabele krzyżowe dla prób złożonych.....	14
Statystyki tabel krzyżowych dla prób złożonych.....	15
Brakujące wartości dla prób złożonych.....	16
Opcje dla prób złożonych.....	16
Statystyki ilorazowe dla prób złożonych.....	16
Statystyki ilorazowe dla prób złożonych.....	17
Brakujące dane w statystykach ilorazowych dla prób złożonych.....	17
Opcje dla prób złożonych.....	18
Ogólny model liniowy dla prób złożonych.....	18
Ogólny model liniowy dla prób złożonych.....	19
Ogólny model liniowy dla prób złożonych: Statystyki.....	19
Próby złożone: Testy hipotez.....	20
Ogólny model liniowy dla prób złożonych: Estymowane średnie.....	20
Ogólny model liniowy dla prób złożonych: Zapisz.....	21

Ogólny model liniowy dla prób złożonych: Opcje.....	21
Dodatkowe właściwości komendy CSGLM.....	21
Regresja logistyczna dla prób złożonych.....	22
Regresja logistyczna dla prób złożonych: Kategoria odniesienia.....	22
Model regresji logistycznej dla prób złożonych.....	23
Regresja logistyczna dla prób złożonych: Statystyki.....	23
Próby złożone: Testy hipotez.....	24
Regresja logistyczna dla prób złożonych: Ilorazy szans.....	24
Regresja logistyczna dla prób złożonych: Zapisz.....	25
Regresja logistyczna prób złożonych: Opcje.....	25
Dodatkowe właściwości komendy CSLOGISTIC.....	26
Regresja porządkowa prób złożonych.....	26
Regresja porządkowa prób złożonych: Prawdopodobieństwa odpowiedzi.....	27
Regresja porządkowa prób złożonych: Model.....	27
Regresja porządkowa prób złożonych: Statystyki.....	28
Próby złożone: Testy hipotez.....	29
Regresja porządkowa prób złożonych: Ilorazy szans.....	29
Regresja porządkowa prób złożonych: Zapisz.....	29
Regresja porządkowa prób złożonych: Opcje.....	30
Dodatkowe właściwości komendy CSORDINAL.....	31
regresja Coxa prób złożonych.....	31
Definiowanie zdarzenia .....	33
Predyktory .....	33
Podgrupy .....	34
Model .....	34
Statystyka .....	34
Wykresy .....	35
Testy hipotezy .....	36
Zapisz .....	36
Eksport .....	37
Wykresy .....	38
Dodatkowe właściwości komendy CSCOXREG.....	39
<b>Uwagi.....</b>	<b>41</b>
Znaki towarowe.....	42
<b>Indeks.....</b>	<b>45</b>

---

# Rozdział 1. Próby złożone

Następujące funkcje związane z próbami złożonymi są dostępne w SPSS Statistics produktu Premium Edition oraz w opcji Complex Samples.

## Wprowadzenie do procedur prób złożonych

---

W procedurach analitycznych w tradycyjnych pakietach oprogramowania niezmiennie obowiązuje założenie, że obserwacje w pliku danych odzwierciedlają prostą próbę losową badanej populacji. To założenie jest kwestionowane przez coraz więcej firm i badaczy, którzy uważają, że pozyskiwanie próby w sposób bardziej ustrukturyzowany jest zarówno bardziej ekonomiczne, jak i dogodne.

Opcja Complex Samples — próby złożone — umożliwia wybór próby zgodnie ze złożonym planem i uwzględnienie specyfikacji planu w analizie danych w celu zapewnienia ważności wyników.

## Własności prób złożonych

Próba złożona może różnić się od zwykłej próby losowej pod wieloma względami. W prostej próbie losowej poszczególne jednostki losowane są wybierane losowo z równym prawdopodobieństwem i bez zastępowania (WOR) bezpośrednio z całej populacji. Dla kontrastu, dana złożona próbka może charakteryzować się częścią lub wszystkimi poniższymi cechami:

**Warstwowania.** Warstwowy dobór próby polega na wyborze prób w sposób niezależny w obrębie nienakładających się na siebie podgrup populacji, nazywanych warstwami. Przykładowe warstwy to grupy socjoekonomiczne, kategorie zawodów, grupy wiekowe lub grupy etniczne. Dzięki podziałowi na warstwy można zapewnić odpowiednią liczebność próby dla interesujących nas podgrup, zwiększyć precyzję ogólnych szacunków i zastosować różne metody losowania prób w poszczególnych warstwach.

**Grupowanie.** Dobór próby poprzez grupowanie polega na wyborze grup jednostek losowanych lub grup. Przykładowe grupy to szkoły, szpitale lub obszary geograficzne, a jednostkami losowanymi mogą być uczniowie, pacjenci lub obywatele. Grupowanie jest często stosowane w planach wieloetapowych i próbach obszarów (geograficznych).

**Wielokrotne etapy.** Wieloetapowy dobór próby polega na wyborze pierwszej próby na podstawie grup. Następnie tworzona jest druga próba poprzez określenie podprób z wybranych grup. Jeśli druga próba opiera się na podgrupach, można dodać trzeci etap do próby. Na przykład na pierwszym etapie ankiety można wybrać próbę miast. Następnie z próby wybranych miast można wybrać próbę gospodarstw domowych. Wreszcie, z próby wybranych gospodarstw domowych można wybrać poszczególne osoby. Kreatory doboru próby i przygotowania analizy umożliwiają określenie trzech etapów planu.

**Nielosowy dobór próby.** Jeśli wybór w sposób losowy jest trudny do uzyskania, jednostki można podzielić na próby w sposób systematyczny (wg stałego przedziału) lub sekwencyjny.

**Nierówne prawdopodobieństwa wyboru.** Podczas wybierania grup, które zawierają nierówną liczbę jednostek, można użyć metody prawdopodobieństwa proporcjonalnego do rozmiaru (PPS), aby prawdopodobieństwo wyboru grupy było równe proporcji jednostek, które zawiera. Podczas doboru próby metodą PPS można również korzystać z bardziej ogólnych schematów ważenia, aby wybrać jednostki.

**Nieograniczony dobór próby.** Nieograniczony dobór prób umożliwia wybieranie jednostek ze zwracaniem (WR). Z tego względu poszczególną jednostkę można wybrać więcej niż jeden raz.

**Wagi losowania.** Wagi prób są obliczane automatycznie podczas określania próby złożonej i idealnie odpowiadają „częstości” reprezentowanej przez poszczególną jednostkę w populacji docelowej. W związku z tym, suma wag w próbie powinna szacować wielkość populacji. W procedurach analizy prób złożonych wymagane są wagi prób w celu prawidłowej analizy próby złożonej. Należy pamiętać, że wagi te powinny być używane w całości w ramach opcji Próby złożone i nie powinny być używane z innymi procedurami analitycznymi poprzez procedurę Ważenie obserwacji, która traktuje wagi jako replikacje przypadków.

## Użycie procedur prób złożonych

Sposób użycia procedur prób złożonych zależy od konkretnych potrzeb. Głównymi typami użytkowników są użytkownicy, którzy:

- Planują i przeprowadzają ankiety według złożonych planów z możliwością analizy próby w późniejszym czasie. Główne narzędzie dla osób tworzących ankiety to [kreator losowania prób](#).
- Analizują uzyskane wcześniej pliki danych zgodnie ze złożonymi planami. Przed użyciem procedur analizy prób złożonych może być konieczne użycie [kreatora przygotowania analizy](#).

Wszyscy użytkownicy muszą wprowadzić informacje o planie do procedur prób złożonych. Informacje te są przechowywane w **pliku planu**, aby można było łatwo użyć ich ponownie.

### Pliki planu

Plik planu zawiera specyfikację próby złożonej. Istnieją dwa typy plików planu:

**Plan losowania.** Specyfikacje podane w kreatorze losowania prób definiują plan próby, który jest używany do wyboru próby złożonej. Plik planu losowania zawiera te specyfikacje. Plik planu losowania zawiera również domyślny plan analizy, w którym wykorzystywane są metody estymacji odpowiednie dla określonego planu próby.

**Plan analizy.** Ten plik planu zawiera informacje potrzebne w procedurach analizy prób złożonych do prawidłowego obliczenia szacunków wariancji dla próby złożonej. Plan zawiera strukturę próby, metody estymacji dla każdego etapu oraz odniesienia do wymaganych zmiennych, takich jak wagi próby. Kreator przygotowania analizy umożliwia tworzenie i edycję planów analiz.

Zapisanie specyfikacji w pliku planu ma wiele korzyści, w tym:

- Osoba tworząca ankietę może określić pierwszy etap wieloetapowego planu losowania i określić jednostki stosowane na pierwszym etapie, zebrać informacje dotyczące jednostek losowanych na drugim etapie, a następnie zmodyfikować plan losowania, aby uwzględnić drugi etap.
- Analityk, który nie ma dostępu do pliku planu losowania, może określić plan analizy i odnieść się do tego planu z każdej procedury analizy prób złożonych.
- Twórca próbek do użytku publicznego na dużą skalę może opublikować plik planu losowania co upraszcza instrukcje dla analityków i eliminuje konieczność definiowania własnych planów analizy przez każdego analityka.

### Dalsza lektura

Aby uzyskać więcej informacji na temat technik doboru próby, zapoznaj się z poniższymi dokumentami:

Cochran, W. G. 1977. *Sampling Techniques*, 3rd ed. New York: John Wiley and Sons.

Kish, L. 1965. *Survey Sampling*. New York: John Wiley and Sons.

Kish, L. 1987. *Statistical Design for Research*. New York: John Wiley and Sons.

Murthy, M. N. 1967. *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.

Särndal, C., B. Swensson, and J. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.

## Próbkowanie z projektu złożonego

Kreator próbkowania prowadzi użytkownika przez kroki związane z tworzeniem, modyfikowaniem lub wykonywaniem pliku planu próbkowania. Przed rozpoczęciem korzystania z kreatora należy zdefiniować dobrze zdefiniowaną populację docelową, listę jednostek próbkowania oraz odpowiedni projekt przykładowy.

## Tworzenie nowego planu losowania

1. Z menu wybierz:

**Analizuj > Próby złożone > Wybierz próbę...**

2. Wybierz opcję **Zaprojektować próbę**, a następnie wybierz nazwę pliku planu, w którym zostanie zapisany plan losowania.
3. Kliknij przycisk **Dalej**, aby kontynuować pracę w kreatorze.
4. Opcjonalnie w kroku Zmienne planu można zdefiniować warstwy, skupienia i wagi zmiennych wejściowych. Po zdefiniowaniu tych informacji, kliknij przycisk **Dalej**.
5. Opcjonalnie w kroku Metoda losowania można wybrać metodę selekcji pozycji.  
W przypadku wybrania opcji **Proporcjonalne Brewera** lub **Proporcjonalne Murthy'ego** można kliknąć przycisk **Zakończ**, aby wylosować próbę. W przeciwnym razie kliknij przycisk **Dalej**, a następnie:
6. W kroku Wielkość próby określ liczbę lub proporcję jednostek do wylosowania.
7. Można teraz kliknąć przycisk **Zakończ**, aby wylosować próbę.

Opcjonalnie na dalszych etapach można:

- Wybierz zmienne wynikowe przeznaczone do zapisania.
- dodać drugi lub trzeci etap do planu;
- Ustaw różne opcje wyboru, w tym etapy, z których mają być losowane próby, wartość początkową generatora liczb losowych oraz to, czy braki danych zdefiniowane przez użytkownika mają być traktowane jako ważne wartości zmiennych planu.
- Wybierz sposób zapisywania danych wynikowych.
- wkleić zaznaczenia jako składnię komend.

## Kreator losowania prób: zmienne projektowe

W tym kroku można wybrać podział na warstwy i zmienne grupowania oraz zdefiniować wejściowe wagi losowania. Można również określić etykietę dla etapu.

**Warstwy według.** Klasyfikacja krzyżowa zmiennych warstwowych definiuje odrębne podpopulacje lub warstwy. Dla każdej warstwy uzyskiwane są oddzielne próby. Aby poprawić precyzję szacunków, jednostki w warstwach powinny być jak najbardziej jednorodne pod względem cech interesujących.

**Grupy.** Zmienne grupowania definiują grupy jednostek obserwacyjnych lub grupy. Grupy są przydatne, gdy bezpośrednie pobieranie próbek jednostek obserwacyjnych z populacji jest kosztowne lub niemożliwe; zamiast tego można pobierać grupy z populacji, a następnie jednostki obserwacyjne z wybranych grup. Użycie grup może jednak wprowadzać korelacje między jednostkami próbkowania, co skutkuje utratą precyzji. Aby zminimalizować ten efekt, jednostki w grupach powinny być jak najbardziej heterogeniczne dla cech interesujących. Aby zaplanować wielostopniową konstrukcję, należy zdefiniować przynajmniej jedną zmienną w grupie. Grupy są również konieczne w przypadku stosowania kilku różnych metod pobierania próbek. Więcej informacji można znaleźć w temacie [“Kreator losowania prób: metoda losowania”](#) na stronie 4.

**Wprowadź wagę próby.** Jeśli bieżąca przykładowa konstrukcja jest częścią większego projektu przykładowego, mogą być dostępne wagi prób z poprzedniego etapu większego projektu. Można określić zmienną numeryczną zawierającą te wagi w pierwszym etapie bieżącego projektu. Wagi przykładowe są obliczane automatycznie dla kolejnych etapów bieżącego projektu.

**Etykieta etapu.** Dla każdego etapu można określić opcjonalną etykietę łańcuchową. Opcja ta jest używana w danych wyjściowych do identyfikowania informacji o etapie.

*Uwaga:* Lista zmiennych źródłowych ma taką samą treść w poszczególnych krokach kreatora. Innymi słowy, zmienne usunięte z listy źródłowej w konkretnym kroku są usuwane z listy we wszystkich krokach. Zmienne zwrócone do listy źródłowej są wyświetlane na liście we wszystkich krokach.

## Drzewa do nawigacji w Kreatorze losowania prób

W każdym kroku Kreatora losowania prób po lewej stronie wyświetlany jest konspekt wszystkich kroków. Można przechodzić między krokami w kreatorze, klikając nazwy aktywnych kroków w konspekcie. Dany krok jest aktywny pod warunkiem, że poprzednie kroki zostały prawidłowo ukończone, tzn. wprowadzono w nich co najmniej wymagane informacje. Więcej informacji o możliwych przyczynach niepoprawności danych zawiera pomoc do poszczególnych kroków.

### Kreator losowania prób: metoda losowania

W tym kroku można określić sposób wybierania obserwacji z aktywnego zbioru danych.

**Metoda.** Elementy sterujące w tej grupie są używane do określania metody wyboru. Niektóre typy losowania pozwalają na losowanie ze zwracaniem (LZZ) lub bez zwracania (LBZ). Więcej informacji można znaleźć w opisach typów. Należy zauważyć, że niektóre prawdopodobieństwa proporcjonalne do wielkości (PPS) są dostępne tylko wtedy, gdy zdefiniowano klastry i że wszystkie typy PPS są dostępne tylko na pierwszym etapie projektowania. Co więcej, metody LZZ są dostępne tylko na ostatnim etapie planu.

- **Losowanie Proste.** Jednostki są wybierane z równym prawdopodobieństwem. Mogą być wybierane ze zwracaniem lub bez zwracania.
- **Proste Systematyczne.** Jednostki są wybierane w stałych odstępach czasu w całej ramce losowania (lub w warstwach, jeśli zostały określone) i wyodrębniane bez zwracania. Jako punkt początkowy wybierana jest losowo wybrana jednostka w obrębie pierwszego przedziału.
- **Proste sekwencyjne.** Jednostki są wybierane kolejno z równym prawdopodobieństwem i bez zwracania.
- **Proporcjonalne (PPS).** Jest to metoda pierwszego etapu, która wybiera jednostki losowo z prawdopodobieństwem proporcjonalnym do wielkości. Ze zwracaniem wybrać dowolne jednostki. Bez zwracania mogą być wybierane tylko skupienia.
- **Proporcjonalne Systematyczne.** Jest to metoda pierwszego etapu, która systematycznie wybiera jednostki z prawdopodobieństwem proporcjonalnym do wielkości. Są one wybierane bez zwracania.
- **Proporcjonalne Sekwencyjne.** Jest to metoda pierwszego etapu, która kolejno wybiera jednostki z prawdopodobieństwem proporcjonalnym do wielkości skupień i bez zwracania.
- **Proporcjonalne Brewera.** Jest to metoda pierwszego etapu, która wybiera dwa skupienia z każdej warstwy z prawdopodobieństwem proporcjonalnym do wielkości skupień i bez zwracania. Aby można było użyć tej metody, należy określić zmienną skupiania.
- **Proporcjonalne Murthy'ego.** Jest to metoda pierwszego etapu, która wybiera dwa skupienia z każdej warstwy z prawdopodobieństwem proporcjonalnym do wielkości skupień i bez zwracania. Aby można było użyć tej metody, należy określić zmienną skupiania.
- **Proporcjonalne Sampforda.** Jest to metoda pierwszego etapu, która wybiera więcej niż dwa skupienia z każdej warstwy z prawdopodobieństwem proporcjonalnym do wielkości skupień i bez zwracania. Jest to rozszerzenie metody Brewera. Aby można było użyć tej metody, należy określić zmienną skupiania.
- **Użyj estymacji LZZ w analizie.** Domyślnie w pliku planu określona jest metoda estymacji, która jest spójna z wybraną metodą próbkowania. Pozwala to na użycie estymacji ze zwracaniem, nawet jeśli metoda losowania implikuje oszacowanie LBZ. Opcja ta jest dostępna wyłącznie w etapie 1.

**Miara wielkości (MW).** Jeśli wybrana jest metoda PPS, należy określić miarę wielkości, która definiuje wielkość każdej jednostki. Wielkości te mogą być określone jawnie w zmiennej lub obliczone z danych. Opcjonalnie można ustawić dolne i górne granice MW, zastępując wszystkie wartości znalezione w zmiennej MW lub obliczone na podstawie danych. Te opcje są dostępne tylko na etapie 1.

### Kreator pobierania próbek: wielkość próby

Ten krok pozwala na określenie liczby lub proporcji jednostek do próbkowania w bieżącym etapie. Wielkość próby może być ustalona lub może się różnić w zależności od warstwy. W celu określenia wielkości próby skupienia wybrane w poprzednich etapach mogą być używane do definiowania warstw.

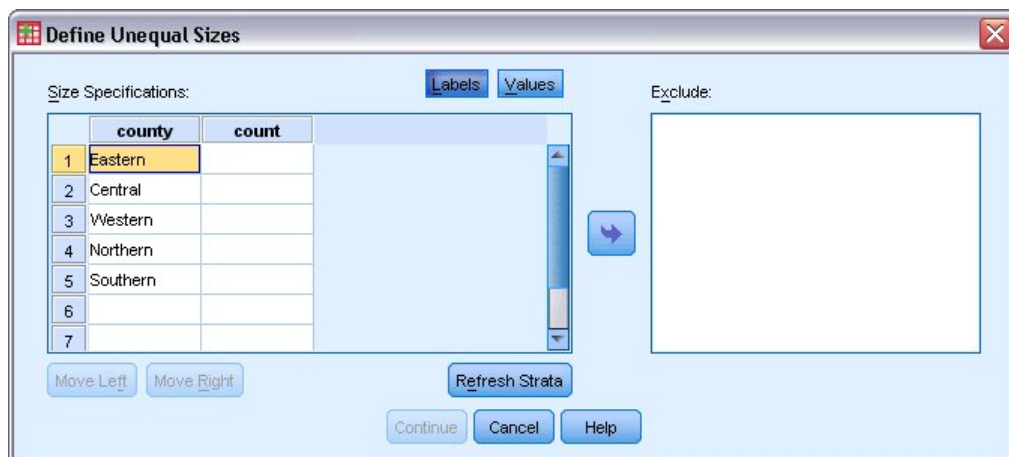
**Jednostki.** Można określić dokładną wielkość próby lub proporcję jednostek do próbkowania.



- **Wartość.** Pojedyncza wartość jest stosowana do wszystkich warstw. Jeśli jako metryka jednostki wybrana jest opcja **Liczebności**, należy wprowadzić dodatnią liczbę całkowitą. Jeśli wybrana jest opcja **Proporcje**, należy wprowadzić wartość nieujemną. O ile nie zostanie próbkowanie z wymianą, wartości proporcji nie powinny być większe niż 1.
- **Nierówne wartości dla warstw.** Umożliwia wprowadzenie wartości wielkości dla poszczególnych warstw za pomocą okna dialogowego definiowania wielkości nierównych wielkości.
- **Odczytaj wartości ze zmiennej.** Umożliwia wybranie zmiennej numerycznej, która zawiera wartości wielkości dla warstw.

If **wybrano Proporcje**, wystąpi opcja ustawienia niższych i wyższych granic liczby próbkowanych jednostek.

## Definiuj nierówne wielkości



Rysunek 1. Okno dialogowe Definiuj nierówne wielkości

Okno dialogowe Definiuj nierówne wielkości umożliwia wprowadzanie wielkości dla poszczególnych warstw.

**Siatka specyfikacji wielkości.** W siatce są wyświetlane klasyfikacje krzyżowe z maksymalnie pięcioma warstwami lub zmiennymi skupień — jedna kombinacja warstwy/skupienia na wiersz. Do zakwalifikowanych zmiennych siatki należą wszystkie zmienne definiujące warstwy z bieżącego i poprzedniego etapu oraz wszystkie zmienne skupienia z poprzednich etapów. Kolejność zmiennych można zmienić w siatce. Można je też przenieść na listę Wyklucz. Wielkości należy wprowadzić w kolumnie znajdującej się najbardziej na prawo. Należy kliknąć opcję **Etykiety** lub **Wartości**, aby przełączyć wyświetlanie etykiet wartości i wartości danych dla zmiennych definiujących warstwy i zmiennych skupień w komórkach siatki. Komórki zawierające wartości bez etykiety zawsze zawierają wartości. Należy kliknąć opcję **Odśwież warstwy**, aby ponownie zapełnić siatkę za pomocą każdej kombinacji wartości danych z etykietami dla zmiennych w siatce.

**Wyklucz.** Aby określić wielkości dla podzbioru kombinacji warstwy/skupienia, należy przenieść co najmniej jedną zmienną na listę Wyklucz. Te zmienne nie są używane do definiowania wielkości prób.

## Kreator losowania prób: zmienne wyników

Ten krok umożliwia wybranie zmiennych, które mają zostać zapisane podczas rysowania próbki.

**Wielkość populacji.** Szacowana liczba jednostek w populacji dla danego etapu. Trzon nazwy zapisanej zmiennej to *PopulationSize\_*.

**Proporcja próby.** Częstotliwość próbkowania na danym etapie. Trzon nazwy zapisanej zmiennej to *SamplingRate\_*.

**Wielkość próby.** Liczba jednostek narysowanych na danym etapie. Trzon nazwy zapisanej zmiennej to *SampleSize\_*.

**Waga próby.** Odwrotność prawdopodobieństw włączenia. Trzon nazwy zapisanej zmiennej to *SampleWeight\_*.

Niektóre zmienne dotyczące etapu są generowane automatycznie. Są to między innymi poniższe elementy:

**Prawdopodobieństwa włączenia.** Proporcja jednostek narysowanych na danym etapie. Trzon nazwy zapisanej zmiennej to *InclusionProbability\_*.

**Waga skumulowana.** Waga skumulowana próby na etapach poprzedzających i bieżącym. Trzon nazwy zapisanej zmiennej to *SampleWeightCumulative\_*.

**Indeks.** Identyfikuje jednostki wybrane wiele razy na danym etapie. Trzon nazwy zapisanej zmiennej to *Index\_*.

*Uwaga:* trzony nazwy zapisanej zmiennej to przyrostek będący liczbą całkowitą, który odzwierciedla numer etapu, na przykład *PopulationSize\_1\_* w przypadku zapisanej wielkości populacji wynoszącej 1.

## Kreator losowania próbek: podsumowanie planu

Jest to ostatni krok na każdym etapie. Udostępnia on podsumowanie specyfikacji projektu próbki na bieżącym etapie. W tym miejscu można przejść do następnego etapu (tworząc go w razie potrzeby) lub ustawić opcje rysowania przykładu.

## Kreator próbkowania: opcje wyboru losowania próbek

Ten krok umożliwia określenie, czy ma być losowana próbka. Można również określić inne opcje próbkowania, takie jak wartość początkowa generatora liczb losowych i obsługa braków wartości.

**Losuj próbkę.** Oprócz wyboru, czy ma być losowana próbka, można również wybrać wykonanie części projektu próbkowania. Etapy muszą być losowane w kolejności, tzn. etap 2 nie może być losowany, chyba że etap 1 został również wylosowany. Podczas edytowania lub wykonywania planu nie można ponownie próbować zablokowanych etapów.

**Wartość początkowa.** Umożliwia wybranie wartości początkowej generatora liczb losowych.

**Uwzględnij brakujące wartości użytkownika.** Określa, czy brakujące wartości użytkownika są poprawne. Jeśli tak, brakujące wartości użytkownika są traktowane jako oddzielna kategoria.

**Dane są już posortowane.** Jeśli ramka próbkowania jest wstępnie określona przez wartości zmiennych warstwowania, ta opcja umożliwia przyspieszenie procesu wyboru.

## Kreator losowania prób: pliki wyjściowe prób losowych

Ten krok umożliwia wybranie miejsca, do którego mają być kierowane losowane próby, zmienne wagi, prawdopodobieństwa łączne i reguły wyboru obserwacji.

**Dane prób.** Te opcje umożliwiają określenie miejsca, w którym zapisywane są dane wyjściowe. Można dodać do aktywnego zbioru danych, zapisać do nowego zbioru danych lub zapisać w zewnętrznym pliku danych IBM SPSS Statistics. Zbiory danych są dostępne podczas bieżącej sesji, ale nie są dostępne w kolejnych sesjach, chyba że zostaną wprost zapisane jako pliki danych. Nazwy zbiorów danych muszą być zgodne z regułami nazewnictwa zmiennych. Jeśli określono plik zewnętrzny lub nowy zbiór danych, zapisywane są zmienne wyjściowe prób i zmienne w aktywnym zbiorze danych dla wybranych obserwacji.

**Prawdopodobieństwa łączne.** Te opcje pozwalają określić, gdzie są zapisywane prawdopodobieństwa łączne. Są one zapisywane w zewnętrznym pliku danych IBM SPSS Statistics. Prawdopodobieństwa łączne są generowane, jeśli wybrano metodę PPS WOR, PPS Brewer, PPS Sampford lub PPS Murthy, a szacowanie WR nie jest określone.

**Reguły wyboru obserwacji.** W przypadku tworzenia próby po jednym etapie na raz, może być przydatne zapisanie reguł wyboru obserwacji w pliku tekstowym. Są one przydatne do konstruowania podramki dla kolejnych etapów.

## Kreator próbkowania: zakończenie

To jest ostatni krok. Można zapisać plik planu i pobrać go teraz lub wkleić wybrane opcje do okna składni.

Podczas wprowadzania zmian w etapach w istniejącym pliku planu można zapisać zmodyfikowany plan w nowym pliku lub nadpisać istniejący plik. Podczas dodawania etapów bez wprowadzania zmian do istniejących etapów, kreator automatycznie nadpisuje istniejący plik planu. Aby zapisać plan w nowym pliku, wybierz opcję **Wklej wygenerowaną składnię kreatora do okna składni** i zmień nazwę pliku w komendach składni.

## Modyfikowanie istniejącego planu losowania

1. Z menu wybierz:

**Analizuj > Próby złożone > Wybierz próbę...**

2. Wybierz opcję **Edytować wzorzec próby**, a następnie wybierz nazwę pliku planu do edycji.

3. Kliknij przycisk **Dalej**, aby kontynuować pracę w kreatorze.

4. Przejrzyj plan losowania w kroku Podsumowanie planu, a następnie kliknij przycisk **Dalej**.

Następne kroki są w dużej mierze takie same, jak w przypadku nowego projektu. Więcej informacji na temat poszczególnych kroków można znaleźć w Pomocy.

5. Przejdź do kroku Zakończ i podaj nową nazwę zmienionego pliku planu lub wybierz opcję nadpisania istniejącego pliku planu.

Opcjonalnie można wykonać następujące czynności:

- Określ etapy, na których przeprowadzono już losowanie.
- Usuń etapy z tego planu.

## Kreator losowania próbek: podsumowanie planu

W tym kroku można przejrzeć plan losowania próbek i wskazać etapy, które już zostały pobrane. Edytując plan, można również usunąć etapy z planu.

**Poprzednio próbkowane etapy.** Jeśli rozszerzona ramka próbkowania nie jest dostępna, konieczne będzie wykonanie wieloetapowego projektowania próbkowania po jednym etapie na raz. Wybierz z listy rozwijanej, które etapy zostały już pobrane. Wszystkie etapy, które zostały wykonane, są zablokowane; nie są dostępne w kroku Opcje wyboru losowania próby i nie mogą być zmieniane podczas edytowania planu.

**Usuń etapy.** Etapy 2 i 3 można usuwać z wieloetapowego projektu.

## Wykonywanie istniejącego planu losowania

1. Wybierz z menu następującą opcję:

**Analizuj > Próby złożone > Wybierz próbę...**

2. Wybierz opcję **Losowanie próby**, a następnie wybierz nazwę pliku planu do wykonania.

3. Kliknij przycisk **Dalej**, aby kontynuować pracę w kreatorze.

4. Przejrzyj plan losowania w kroku Podsumowanie planu, a następnie kliknij przycisk **Dalej**.

5. Podczas wykonywania planu losowania pomijane są poszczególne kroki zawierające informacje o etapach. W dowolnym momencie można teraz przejść do kroku Zakończ.

Opcjonalnie można określić etapy, na których przeprowadzono już losowanie.

## Dodatkowe funkcje komend CSPLAN i CSSELECT

Język składni komend umożliwia również:

- Określanie niestandardowych nazw zmiennych wynikowych.

- Sterowanie wynikami w oknie raportów. Na przykład można ukryć etapowe podsumowanie planu wyświetlane po utworzeniu lub zmodyfikowaniu losowania, ukryć podsumowanie rozkładu losowanych obserwacji według warstw widoczne po wykonaniu planu losowania, a także zażądać podsumowania przetwarzania obserwacji.
- Wybierz podzbiór zmiennych w aktywnym zestawie danych, który ma zostać zapisany w zewnętrznym pliku próby lub w innym zbiorze danych.

Pełne informacje na temat składni znajdują się w podręczniku *Command Syntax Reference*.

## Przygotowywanie próbki złożonej do analizy

Kreator przygotowania analizy prowadzi użytkownika przez kroki związane z tworzeniem lub modyfikowaniem planu analizy do użycia z różnymi procedurami analizy próbek złożonych. Przed rozpoczęciem korzystania z kreatora należy pobrać próbkę pobraną zgodnie ze złożonym projektem.

Utworzenie nowego planu jest najbardziej przydatne w sytuacji, gdy użytkownik nie ma dostępu do pliku planu losowania używanego do pobierania próbki (przypomnijmy, że plan losowania zawiera domyślny plan analizy). Jeśli użytkownik ma dostęp do pliku planu losowania używanego do pobierania próbki, może użyć domyślnego planu analizy zawartego w pliku planu losowania lub zastąpić domyślne specyfikacje analizy i zapisać zmiany w nowym pliku.

### Tworzenie nowego planu analizy

1. Z menu wybierz:

**Analiza > Próby złożone > Przygotuj do analizy...**

2. Wybierz opcję **Utwórz plik planu**, a następnie wybierz nazwę pliku planu, w którym zostanie zapisany plan analizy.
3. Kliknij przycisk **Dalej**, aby kontynuować pracę w kreatorze.
4. Określ zmienną zawierającą wagi prób na etapie Zmienne planu, opcjonalnie definiując warstwy i skupienia.
5. Można teraz kliknąć przycisk **Zakończ**, aby zapisać plan.

Opcjonalnie na dalszych etapach można:

- wybrać metodę estymacji błędów standardowych na etapie Metoda estymacji;
- określić liczbę losowanych jednostek lub prawdopodobieństwo włączenia jednostki na etapie Rozmiar;
- dodać drugi lub trzeci etap do planu;
- wkleić zaznaczenia jako składnię komend.

### Kreator przygotowania analizy: zmienne projektowe

Ten krok pozwala na zidentyfikowanie zmiennych warstw i grupowania oraz definiowanie wag próby. Można również podać etykietę dla etapu.

**Warstwa.** Klasyfikacja krzyżowa zmiennych warstwowych definiuje odrębne podpopulacje lub warstwy. Łączna liczba próbek reprezentuje kombinację prób niezależnych z każdej warstwy.

**Grupy.** Zmienne grupowania definiują grupy jednostek obserwacyjnych lub grupy. Próby wylosowane w wielu etapach wybierają grupy we wcześniejszych etapach, a następnie jednostki podprób z wybranych grup. Analizując plik danych otrzymany przez grupowanie prób z wymianą, należy uwzględnić indeks duplikowania jako zmienną grupowania.

**Waga próby.** Wagi próby należy podać w pierwszym etapie. Wagi próby są obliczane automatycznie dla kolejnych etapów bieżącego projektu.

**Etykieta etapu.** Dla każdego etapu można określić opcjonalną etykietę łańcuchową. Opcja ta jest używana w danych wyjściowych do identyfikowania informacji o etapie.

*Uwaga:* Lista zmiennych źródłowych ma taką samą treść we wszystkich krokach kreatora. Innymi słowy, zmienne usunięte z listy źródłowej w konkretnym kroku są usuwane z listy we wszystkich krokach. Zmienne zwrócone do listy źródłowej są wyświetlane we wszystkich krokach.

## Drzewa do nawigacji w Kreatorze przygotowania analiz

W każdym kroku Kreatora przygotowywania analiz po lewej stronie wyświetlany jest konspekt wszystkich kroków. Można przechodzić między krokami w kreatorze, klikając nazwy aktywnych kroków w konspekcie. Dany krok jest aktywny pod warunkiem, że poprzednie kroki zostały prawidłowo ukończone, tzn. wprowadzono w nich co najmniej wymagane informacje. Więcej informacji o możliwych przyczynach niepoprawności danych zawiera Pomoc do poszczególnych kroków.

## Kreator przygotowania analizy: metoda szacowania

W tym kroku można określić metodę szacowania dla etapu.

**WR (losowanie z wymianą).** Oszacowanie WR nie obejmuje korekty losowania ze skończonej populacji (FPC) podczas szacowania wariancji w ramach złożonego planu losowania. Można włączyć lub wyłączyć korektę FPC podczas szacowania wariancji w ramach prostej próby losowej (SRS).

Nieuwzględnianie korekty FPC dla oszacowania wariancji SRS jest zalecane w przypadku, gdy wagi analizy zostały skalowane tak, aby się nie sumowały do wielkości populacji. Oszacowanie wariancji SRS jest używane w statystykach obliczeniowych, takich jak efekt planu. Oszacowanie WR można określić tylko w końcowym etapie projektowania; kreator nie zezwoli na dodanie kolejnego etapu, jeśli wybrano oszacowanie WR.

**Równe WOR (równe prawdopodobieństwo losowania bez wymiany).** Oszacowanie równego WOR obejmuje korektę populacji skończonej i zakłada, że jednostki są losowane z równym prawdopodobieństwem. Oszacowanie równego WOR można określić na dowolnym etapie projektowania.

**Nierówne WOR (różne prawdopodobieństwo losowania bez wymiany).** Oprócz używania korekty populacji skończonej, konta z oszacowaniem nierównego WOR dla jednostek losowania (zwykle skupień) wybranych z różnym prawdopodobieństwem. Ta metoda szacowania jest dostępna tylko w pierwszym etapie.

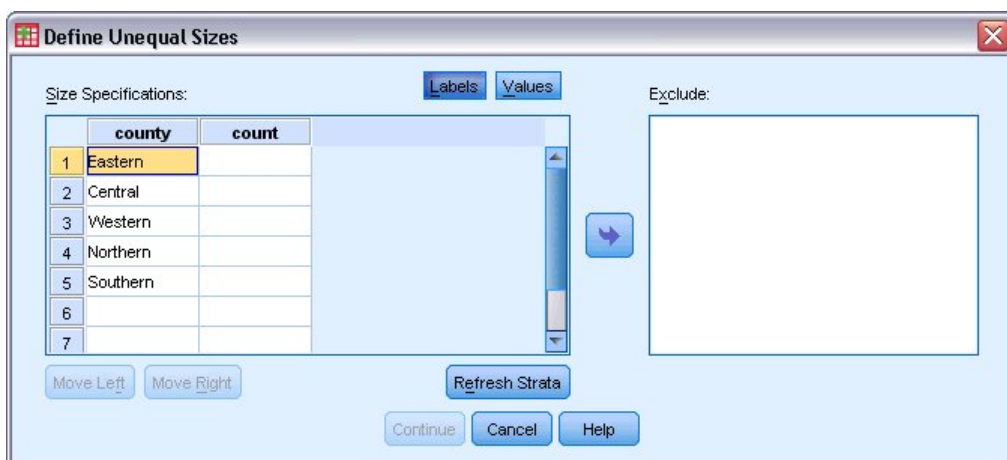
## Kreator przygotowania analizy: wielkość

Ten krok jest używany do określania prawdopodobieństw włączenia lub wielkości populacji dla bieżącego etapu. Wielkości mogą być stałe lub mogą się różnić w zależności od warstwy. W celu określenia wielkości, skupienia określone w poprzednich etapach mogą być używane do definiowania warstw. Należy zauważyć, że ten krok jest wymagany tylko wtedy, gdy jako metoda oszacowania wybrana jest wartość WOR.

**Jednostki.** Możliwe jest określenie dokładnej wielkości populacji lub prawdopodobieństw, z którymi jednostki są próbkowane.

- **Wartość.** Pojedyncza wartość jest stosowana do wszystkich warstw. Jeśli jako metryka jednostki wybrana jest opcja **Wielkości populacji**, należy wprowadzić nieujemną liczbę całkowitą. Jeśli wybrana jest opcja **Prawdopodobieństwa włączenia**, należy wprowadzić wartość z zakresu od 0 do 1 włącznie.
- **Nierówne wartości dla warstw.** Umożliwia wprowadzenie wartości wielkości dla poszczególnych warstw za pomocą okna dialogowego definiowania wielkości nierównych wielkości.
- **Odczytaj wartości ze zmiennej.** Umożliwia wybranie zmiennej numerycznej, która zawiera wartości wielkości dla warstw.

## Definiuj nierówne wielkości



Rysunek 2. Okno dialogowe Definiuj nierówne wielkości

Okno dialogowe Definiuj nierówne wielkości umożliwia wprowadzanie wielkości dla poszczególnych warstw.

**Siatka specyfikacji wielkości.** W siatce są wyświetlane klasyfikacje krzyżowe z maksymalnie pięcioma warstwami lub zmiennymi skupień — jedna kombinacja warstwy/skupienia na wiersz. Do zakwalifikowanych zmiennych siatki należą wszystkie zmienne definiujące warstwy z bieżącego i poprzedniego etapu oraz wszystkie zmienne skupienia z poprzednich etapów. Kolejność zmiennych można zmienić w siatce. Można je też przenieść na listę Wyklucz. Wielkości należy wprowadzić w kolumnie znajdującej się najbardziej na prawo. Należy kliknąć opcję **Etykiety** lub **Wartości**, aby przełączyć wyświetlanie etykiet wartości i wartości danych dla zmiennych definiujących warstwy i zmiennych skupień w komórkach siatki. Komórki zawierające wartości bez etykiety zawsze zawierają wartości. Należy kliknąć opcję **Odśwież warstwy**, aby ponownie zapełnić siatkę za pomocą każdej kombinacji wartości danych z etykietami dla zmiennych w siatce.

**Wyklucz.** Aby określić wielkości dla podzbioru kombinacji warstwy/skupienia, należy przenieść co najmniej jedną zmienną na listę Wyklucz. Te zmienne nie są używane do definiowania wielkości prób.

## Kreator przygotowania analizy: podsumowanie planu

Jest to ostatni krok na każdym etapie. Udostępnia on podsumowanie specyfikacji projektu analizy na bieżącym etapie. W tym miejscu można przejść do następnego etapu (w razie potrzeby utworzyć go) lub zapisać specyfikacje analizy.

Jeśli nie można dodać kolejnego etapu, jest to prawdopodobnie spowodowane jedną z poniższych przyczyn:

- W kroku Zmienne projektu nie określono zmiennej klastra.
- W kroku Metoda estymacji wybrano oszacowanie WR.
- Jest to trzeci etap analizy, a kreator obsługuje maksymalnie trzy etapy.

## Kreator przygotowania analizy: zakończenie

To jest ostatni krok. Plik planu można zapisać teraz lub wkleić wybrane opcje do okna składni.

Podczas wprowadzania zmian w etapach w istniejącym pliku planu można zapisać zmodyfikowany plan w nowym pliku lub nadpisać istniejący plik. Podczas dodawania etapów bez wprowadzania zmian do istniejących etapów, kreator automatycznie nadpisuje istniejący plik planu. Aby zapisać plan w nowym pliku, wybierz opcję **Wklej wygenerowaną składnię kreatora do okna składni** i zmień nazwę pliku w komendach składni.

## Modyfikowanie istniejącego planu analizy

1. Z menu wybierz:

**Analiza > Próby złożone > Przygotuj do analizy...**

2. Wybierz opcję **Edytuj plik planu**, a następnie wybierz nazwę pliku planu, w którym zostanie zapisany plan analizy.

3. Kliknij przycisk **Dalej**, aby kontynuować pracę w kreatorze.

4. Przejrzyj plan analizy w kroku Podsumowanie planu, a następnie kliknij przycisk **Dalej**.

Następne kroki są w dużej mierze takie same, jak w przypadku nowego projektu. Aby uzyskać więcej informacji, patrz Pomoc do poszczególnych kroków.

5. Przejdź do kroku Zakończ i podaj nową nazwę zmienionego pliku planu lub wybierz opcję nadpisania istniejącego pliku planu.

Opcjonalnie można usunąć etapy z planu.

## Kreator przygotowania analizy: podsumowanie planu

Ten krok pozwala na przejrzanie planu analizy i usunięcie etapów z planu.

**Usuń etapy.** Etapy 2 i 3 można usuwać z wieloetapowego projektu. Ponieważ plan musi mieć przynajmniej jeden etap, etap 1 można edytować, ale nie można go usunąć z projektu.

## Plan prób złożonych

---

Procedury analizy dla prób złożonych wymagają specyfikacji analizy z pliku analizy lub pliku planu prób w celu udostępnienia poprawnych wyników.

**Zaplanuj.** Określ ścieżkę pliku analizy lub pliku planu prób.

**Prawdopodobieństwa łączne.** Aby użyć oszacowania nierównego WOR dla skupień wykreślanych za pomocą metody PPS WOR, należy określić osobny plik lub otwarty zbiór danych zawierający prawdopodobieństwa łączne. Ten plik lub zbiór danych jest tworzony przy użyciu kreatora próbkowania podczas próbkowania.

## Częstości dla prób złożonych

---

Procedura Częstości dla prób złożonych tworzy tabele częstości dla wybranych zmiennych i wyświetla statystyki jednej zmiennej. Opcjonalnie można wyświetlić statystyki według podgrup zdefiniowanych przy użyciu co najmniej jednej zmiennej jakościowej.

**Przykład.** Procedura Częstości dla prób złożonych pozwala uzyskać statystyki tabelaryczne jednej zmiennej dotyczące spożycia witamin przez obywateli Stanów Zjednoczonych na podstawie wyników ankiety National Health Interview Survey (NHIS) oraz przy użyciu odpowiedniego planu analizy danych publicznych.

**Statystyki.** Ta procedura generuje oszacowania dotyczące wielkości populacji komórek i wartości procentowych tabeli, a także błędy standardowe, przedziały ufności, współczynniki zmienności, efekty projektowe, pierwiastki kwadratowe efektów projektowych, wartości skumulowane i nieważone liczebności dla każdego oszacowania. Dodatkowo statystyki chi-kwadrat i ilorazu wiarygodności są obliczane dla testu o jednakowych proporcjach komórek.

Uwagi dotyczące danych procedury Częstości dla prób złożonych

**Dane.** Zmienne, dla których tworzone są tabele częstości, powinny być jakościowe. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym Plan prób złożonych.

Uzyskiwanie częstości dla prób złożonych

1. Z menu wybierz:

**Analiza > Przykłady złożone > Częstości...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Częstości dla prób złożonych wybierz co najmniej jedną zmienną częstości.

Opcjonalnie można określić zmienne definiujące podpopulacje. Statystyki dla każdej podpopulacji są obliczane oddzielnie.

## Statystyki częstości dla prób złożonych

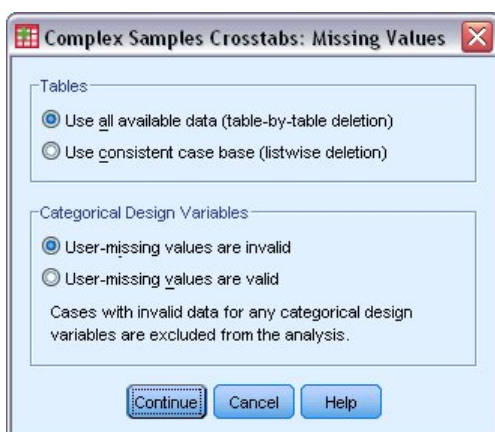
**Komórki.** Ta grupa umożliwia żądanie oszacowania wielkości populacji komórek oraz wartości procentowych tabel.

**Statystyki.** Ta grupa tworzy statystyki powiązane z wielkością populacji lub wartościami procentowymi tabel.

- **Błąd standardowy.** Błąd standardowy oszacowania.
- **Przedział ufności.** Przedział ufności dla oszacowania z uwzględnieniem określonego poziomu.
- **Współczynnik zmienności.** Stosunek błędu standardowego oszacowania do oszacowania.
- **Nieważona liczebność.** Liczba jednostek użytych do obliczenia wartości szacunkowej.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Wartości skumulowane.** Skumulowana wartość szacunkowa dla każdej wartości zmiennej.

**Test jednakowych proporcji komórek.** Powoduje to wykonanie testów chi-kwadrat i ilorazu wiarygodności hipotezy, że kategorie zmiennej mają równe częstości. Dla każdej zmiennej wykonywane są osobne testy.

## Brakujące wartości dla prób złożonych



Rysunek 3. Okno dialogowe Brakujące wartości

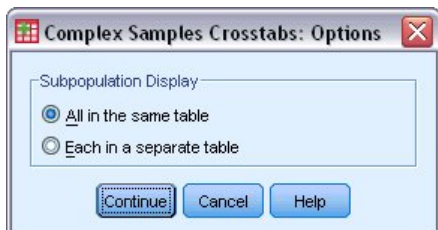
**Tabele.** Ta grupa określa, które próby, które są używane w analizie.



- **Użyj wszystkich dostępnych danych.** Brakujące wartości są określane dla poszczególnych tabel. W związku z tym obserwacje używane do obliczania statystyk mogą być różne w różnych tabelach częstości lub tabelach krzyżowych.
- **Użyj spójnej podstawy danych.** Brakujące wartości są określane dla wszystkich zmiennych. W związku z tym obserwacje używane do obliczania statystyk są spójne w tabelach.

**Zmienne projektu jakościowego.** Ta grupa określa, czy braki danych zdefiniowane przez użytkownika są poprawne.

## Opcje dla prób złożonych



Rysunek 4. Okno dialogowe Opcje

**Wyświetlanie podpopulacji.** Istnieje możliwość określenia, czy podpopulacje mają być wyświetlane w tej samej tabeli, czy w oddzielnych tabelach.

## Próby złożone - Statystyki opisowe

Procedura Próby złożone - Statystyki opisowe wyświetla statystyki podsumowujące dla kilku zmiennych. Opcjonalnie można zażądać statystyk według podgrup zdefiniowanych przez jedną lub więcej zmiennych jakościowych.

**Przykład.** Za pomocą procedury Próby złożone - Statystyki opisowe można uzyskać statystyki opisowe jednej zmiennej dla poziomów aktywności obywateli Stanów Zjednoczonych na podstawie wyników National Health Interview Survey (NHIS) oraz z odpowiednim planem analizy dla tych publicznie dostępnych danych.

**Statystyki.** W ramach tej procedury tworzone są średnie i sumy, a także testy  $T$ , standardowe błędy, przedziały ufności, współczynniki zmienności, liczebności nieważone, wielkości populacji, efekty projektowe i pierwiastki kwadratowe efektów projektowych dla każdego oszacowania.

Uwagi dotyczące prób złożonych statystyk opisowych

**Dane.** Miary powinny być zmiennymi ilościowymi. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym Plan prób złożonych.

Uzyskiwanie prób złożonych statystyk opisowych

1. Z menu wybierz:

**Analizuj > Próby złożone > Statystyki opisowe...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Próby złożone - Statystyki opisowe wybierz przynajmniej jedną zmienną miary.

Opcjonalnie można określić zmienne definiujące podpopulacje. Statystyki są obliczane dla każdej podpopulacji oddzielnie.

## Statystyki opisowe dla prób złożonych

**Podsumowania.** Ta grupa umożliwia żądanie oszacowania średnich i sum zmiennych miar. Ponadto istnieje możliwość żądania testów  $t$  oszacowań dla określonej wartości.

**Statystyki.** Ta grupa tworzy statystyki powiązane ze średnią lub sumą.

- **Błąd standardowy.** Błąd standardowy oszacowania.
- **Przedział ufności.** Przedział ufności dla oszacowania z uwzględnieniem określonego poziomu.
- **Współczynnik zmienności.** Stosunek błędu standardowego oszacowania do oszacowania.
- **Nieważona liczebność.** Liczba jednostek użytych do obliczenia wartości szacunkowej.
- **Wielkość populacji.** Szacowana liczba jednostek w populacji.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.

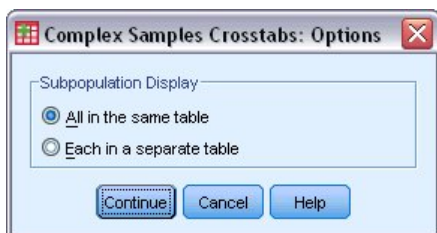
## Brakujące wartości w statystykach opisowych dla prób złożonych

**Statystyki dla zmiennych pomiarów.** Ta grupa określa, które próby, które są używane w analizie.

- **Użyj wszystkich dostępnych danych.** Brakujące wartości są określane osobno dla każdej zmiennej, a zatem próby używane do obliczania statystyk mogą być różne w różnych zmiennych pomiarów.
- **Zapewnienie spójnej podstawy danych.** Brakujące wartości są określane dla wszystkich zmiennych, a zatem przypadki używane do obliczania statystyk są spójne.

**Zmienne projektu jakościowego.** Ta grupa określa, czy braki danych zdefiniowane przez użytkownika są poprawne.

## Opcje dla prób złożonych



Rysunek 5. Okno dialogowe Opcje

**Wyświetlanie podpopulacji.** Istnieje możliwość określenia, czy podpopulacje mają być wyświetlane w tej samej tabeli, czy w oddzielnych tabelach.

## Tabele krzyżowe dla prób złożonych

Procedura Tabele krzyżowe dla prób złożonych powoduje utworzenie tabel krzyżowych dla par wybranych zmiennych i wyświetla statystyki dwukierunkowe. Opcjonalnie można zażądać statystyk według podgrup zdefiniowanych przez jedną lub więcej zmiennych jakościowych.

**Przykład.** Za pomocą procedury Tabele krzyżowe dla prób złożonych można uzyskać statystyki klasyfikacji krzyżowej dla częstotliwości palenia według użycia witamin dla obywateli USA w oparciu o wyniki National Health Interview Survey (NHIS) oraz z odpowiednim planem analizy dla tych danych publicznych.

**Statystyki.** Ta procedura generuje oszacowania wielkości populacji komórek i wartości procentowe wierszy, kolumn i tabel, a także standardowe błędy, przedziały ufności, współczynniki zmienności, oczekiwane wartości, efekty projektowe, pierwiastki kwadratowe efektów projektowych, reszty,

skorygowane reszty i nieważone liczebności dla każdego oszacowania. Iloraz szans, względne ryzyko i różnica ryzyka są obliczane dla tabel 2x2. Dodatkowo statystyki Pearsona i ilorazu wiarygodności są obliczane dla testu niezależności zmiennych w wierszach i kolumnach.

Uwagi na temat danych tabel krzyżowych prób złożonych

**Dane.** Zmienne wiersza i kolumny powinny być jakościowe. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym Plan prób złożonych.

Uzyskiwanie tabel krzyżowych dla prób złożonych

1. Wybierz z menu następującą opcję:

**Analizuj > Próby złożone > Tabele krzyżowe...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Tabele krzyżowe dla prób złożonych wybierz przynajmniej jedną zmienną wierszową i jedną zmienną kolumnową.

Opcjonalnie można określić zmienne definiujące podpopulacje. Statystyki są obliczane dla każdej podpopulacji oddzielnie.

## Statystyki tabel krzyżowych dla prób złożonych

**Komórki.** Ta grupa umożliwia żądanie oszacowania wielkości populacji komórek oraz wartości procentowych wierszy, kolumn i tabel.

**Statystyki.** Ta grupa tworzy statystyki powiązane z wielkością populacji i procentami w wierszach, kolumnach i tabelach.

- **Błąd standardowy.** Błąd standardowy oszacowania.
- **Przedział ufności.** Przedział ufności dla oszacowania z uwzględnieniem określonego poziomu.
- **Współczynnik zmienności.** Stosunek błędu standardowego oszacowania do oszacowania.
- **Wartości oczekiwane.** Wartość oczekiwania oszacowania, zgodnie z hipotezą niezależności tabel krzyżowych dla prób złożonych.
- **Nieważona liczebność.** Liczba jednostek użytych do obliczenia wartości szacunkowej.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Reszty.** Oczekiwana wartość to liczba obserwacji, które należy oczekiwać w komórce, gdy między tymi dwiema zmiennymi nie ma relacji. Dodatnia reszta wskazuje, że w komórce występuje więcej obserwacji niż w przypadku gdyby zmienne wierszowa i kolumnowa były niezależne.
- **Reszty skorygowane.** Reszta dla komórki (obserwowana minus oczekiwana wartość) podzielona przez oszacowanie jej błędu standardowego. Otrzymana standaryzowana reszta jest wyrażana w jednostkach odchylenia standardowego powyżej lub poniżej średniej.

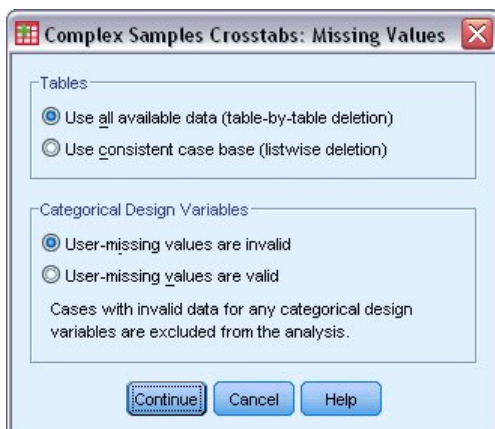
**Podsumowania dla tabel 2 na 2.** Ta grupa tworzy statystyki dla tabel, w których zmienna wierszowa i zmienna kolumnowa mają dwie kategorie. Każdy z nich jest miarą siły związku między obecnością czynnika a wystąpieniem zdarzenia.

- **Iloraz szans.** Iloraz szans może być stosowany jako oszacowanie względnego ryzyka, gdy wystąpienie czynnika jest rzadkie.

- **Względne ryzyko.** Stosunek ryzyka wystąpienia zdarzenia w obecności czynnika do ryzyka zdarzenia w przypadku braku czynnika.
- **Różnica ryzyka.** Różnica między ryzykiem wystąpienia zdarzenia w obecności czynnika a ryzykiem wystąpienia zdarzenia w przypadku braku czynnika.

**Test niezależności wierszy i kolumn.** Tworzy testy chi-kwadrat i ilorazu wiarygodności hipotezy dotyczącej niezależności wiersza i kolumny. Dla każdej pary zmiennych wykonywane są osobne testy.

## Brakujące wartości dla prób złożonych



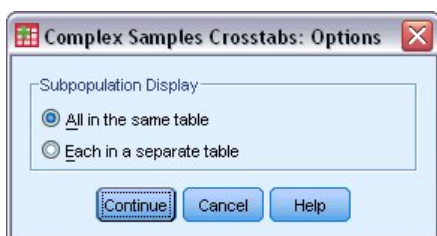
Rysunek 6. Okno dialogowe Brakujące wartości

**Tabele.** Ta grupa określa, które próby, które są używane w analizie.

- **Użyj wszystkich dostępnych danych.** Brakujące wartości są określane dla poszczególnych tabel. W związku z tym obserwacje używane do obliczania statystyk mogą być różne w różnych tabelach częstości lub tabelach krzyżowych.
- **Użyj spójnej podstawy danych.** Brakujące wartości są określane dla wszystkich zmiennych. W związku z tym obserwacje używane do obliczania statystyk są spójne w tabelach.

**Zmienne projektu jakościowego.** Ta grupa określa, czy braki danych zdefiniowane przez użytkownika są poprawne.

## Opcje dla prób złożonych



Rysunek 7. Okno dialogowe Opcje

**Wyświetlanie podpopulacji.** Istnieje możliwość określenia, czy podpopulacje mają być wyświetlane w tej samej tabeli, czy w oddzielnych tabelach.

## Statystyki ilorazowe dla prób złożonych

Procedura Statystyki ilorazowe dla prób złożonych wyświetla statystyki podsumowujące jednej zmiennej dla ilorazu zmiennych. Opcjonalnie można wyświetlić statystyki według podgrup, zdefiniowanych przy użyciu co najmniej jednej zmiennej jakościowej.

**Przykład.** Procedura Statystyki ilorazowe dla prób złożonych pozwala uzyskać statystyki opisowe dla ilorazu bieżącej wartości nieruchomości i ostatniej oszacowanej wartości na podstawie wyników ankiety przeprowadzonej w całym kraju zgodnie ze złożonym planem oraz z odpowiednim planem analizy danych.

**Statystyki.** W ramach tej procedury tworzone są ilorazy, a także testy  $T$ , standardowe błędy, przedziały ufności, współczynniki zmienności, liczebności nieważone, wielkości populacji, efekty projektowe i pierwiastki kwadratowe efektów projektowych dla każdego oszacowania.

Uwagi na temat danych procedury Statystyki ilorazowe dla prób złożonych

**Dane.** Licznikami i mianownikami powinny być zmienne ilościowe o wartości dodatniej. Zmienne podpopulacji mogą być tańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym Plan prób złożonych.

Uzyskiwanie statystyk ilorazowych dla prób złożonych

1. Z menu wybierz:

**Analiza > Próby złożone > Ilorazy...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Statystyki ilorazowe dla prób złożonych wybierz co najmniej jedną zmienną licznika i zmienną mianownika.

Opcjonalnie można określić zmienne, aby zdefiniować podgrupy, dla których generowane są statystyki.

## Statystyki ilorazowe dla prób złożonych

**Statystyki.** Ta grupa tworzy statystyki powiązane z oszacowaniem ilorazu.

- **Błąd standardowy.** Błąd standardowy oszacowania.
- **Przedział ufności.** Przedział ufności dla oszacowania z uwzględnieniem określonego poziomu.
- **Współczynnik zmienności.** Stosunek błędu standardowego oszacowania do oszacowania.
- **Nieważona liczebność.** Liczba jednostek użytych do obliczenia wartości szacunkowej.
- **Wielkość populacji.** Szacowana liczba jednostek w populacji.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.

**Test t.** Istnieje możliwość żądania testów  $t$  oszacowań dla określonej wartości.

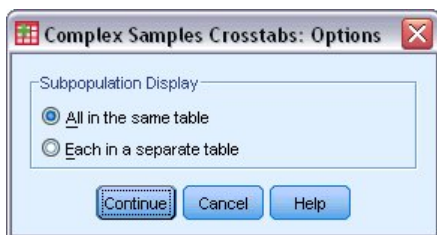
## Brakujące dane w statystykach ilorazowych dla prób złożonych

**Statystyki ilorazowe.** Ta grupa określa, które obserwacje są używane w analizie.

- **Użyj wszystkich dostępnych danych.** Brakujące dane są określane dla poszczególnych statystyk ilorazowych. W związku z tym obserwacje używane do obliczania statystyk mogą być różne w różnych parach licznika i mianownika.
- **Zapewnienie spójnej podstawy danych.** Brakujące dane są określane dla wszystkich zmiennych. W związku z tym obserwacje używane do obliczania statystyk są spójne.

**Zmienne projektu jakościowego.** Ta grupa określa, czy braki danych zdefiniowane przez użytkownika są poprawne.

## Opcje dla prób złożonych



Rysunek 8. Okno dialogowe Opcje

**Wyświetlanie podpopulacji.** Istnieje możliwość określenia, czy podpopulacje mają być wyświetlane w tej samej tabeli, czy w oddzielnych tabelach.

## Ogólny model liniowy dla prób złożonych

Procedura Ogólny model liniowy dla prób złożonych (CSGLM) wykonuje analizę regresji liniowej oraz analizę wariancji i kowariancji próbek pobranych za pomocą skomplikowanych metod próbkowania. Opcjonalnie można zlecić analizy podpopulacji.

**Przykład.** Sieć sklepów spożywczych przeprowadziła wśród grupy klientów złożoną ankietę dotyczącą nawyków zakupowych. Dzięki wynikom ankiety oraz danym na temat kwoty wydanej przez każdego klienta w poprzednim miesiącu, sklep chce dowiedzieć się, czy częstotliwość, z jaką klient robi zakupy w sklepie jest powiązana z miesięcznymi wydatkami. W badaniach brana jest pod uwagę płeć klienta oraz uwzględniony jest plan losowania.

**Statystyki.** Procedura tworzy oszacowania, błędy standardowe, przedziały ufności, testy  $t$ , efekty planu i pierwiastki kwadratowe efektów planu dla parametrów modelu, jak również korelacje i kowariancje między oszacowaniami parametrów. Dostępne są również miary dopasowania modelu oraz statystyki opisowe dla zmiennych zależnych i niezależnych. Ponadto można zażądać szacowanych średnich brzegowych dla poziomów czynników modelu i interakcji czynników.

Ogólny model liniowy dla prób złożonych: Wymagania dotyczące danych

**Dane.** Zmienna zależna jest ilościowa. Czynniki są typu jakościowego. Współzmiennie są zmiennymi ilościowymi powiązane ze zmienną zależną. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje w pliku danych reprezentują próbkę z projektu złożonego, która ma być analizowana zgodnie z ustawieniami w pliku wybranym w [oknie dialogowym Plan prób złożonych](#).

Otrzymywanie ogólnego modelu liniowego dla prób złożonych

1. Z menu wybierz:

**Analiza > Próby złożone > Ogólny model liniowy...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Ogólny model liniowy dla prób złożonych wybierz zmienną zależną.

Opcjonalnie można wykonać następujące czynności:

- Wybrać zmienne dla czynników i współzmiennych odpowiednio do danych.
- Określić zmienną, która ma definiować podpopulację. Analiza jest wykonywana tylko dla wybranej kategorii zmiennej podpopulacji.

## Ogólny model liniowy dla prób złożonych

**Określenie efektów modelu.** Domyślnie procedura buduje model efektów głównych na podstawie czynników i współzmiennych określonych w głównym oknie dialogowym. Można też zbudować model niestandardowy zawierający efekty interakcji i składniki zagnieżdżone.

Składniki niezagnieżdżone

Dla wybranych czynników i współzmiennych:

**Interakcja.** Dla wszystkich wybranych zmiennych tworzy składnik interakcji najwyższego rzędu.

**Efekty główne.** Dla każdej wybranej zmiennej tworzy składnik efektów głównych.

**Wszystkie 2 rzędu.** Tworzy wszystkie możliwe interakcje drugiego rzędu wybranych zmiennych.

**Wszystkie 3 rzędu.** Tworzy wszystkie możliwe interakcje trzeciego rzędu wybranych zmiennych.

**Wszystkie 4 rzędu.** Tworzy wszystkie możliwe interakcje czwartego rzędu wybranych zmiennych.

**Wszystkie 5 rzędu.** Tworzy wszystkie możliwe interakcje piątego rzędu wybranych zmiennych.

Składniki zagnieżdżone

W tej procedurze można zbudować składniki zagnieżdżone dla modelu. Składniki zagnieżdżone są przydatne do modelowania efektu czynników lub współzmiennych, których wartości nie wchodzi w interakcje z poziomami innego czynnika. Na przykład sieć sklepów spożywczych może analizować zwyczaje zakupowe swoich klientów w kilku sklepach. Ponieważ każdy klient bywa regularnie tylko w jednym z tych sklepów, efekt *Klient* jest **zagnieżdżony w** efekcie *Lokalizacja sklepu*.

Ponadto można uwzględnić efekty interakcji, takie jak składniki wielomianowe z tą samą współzmienną, lub dodać wiele poziomów zagnieżdżenia do składnika zagnieżdżonego.

**Ograniczenia.** W odniesieniu do składników zagnieżdżonych obowiązują następujące ograniczenia:

- Wszystkie czynniki w interakcji muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A^*A$ .
- Wszystkie czynniki w efekcie zagnieżdżonym muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A(A)$ .
- Efekt nie może być zagnieżdżony w obrębie współzmiennnej. A zatem, jeśli  $A$  jest czynnikiem, a  $X$  jest współzmienną, to określenie  $A(X)$  jest niedozwolone.

**Wyraz wolny.** Wyraz wolny jest zwykle uwzględniany w modelu. Jeśli można założyć, że dane przechodzą przez początek układu współrzędnych, wyraz wolny można wyłączyć z modelu. Nawet jeśli model zawiera wyraz wolny, można pominąć związane z nim statystyki.

## Ogólny model liniowy dla prób złożonych: Statystyki

**Parametry modelu.** Ta grupa umożliwia sterowanie wyświetlaniem statystyk związanych z parametrami modelu.

- **Oszacowanie.** Wyświetla oszacowania współczynników.
- **Błąd standardowy.** Wyświetla błąd standardowy dla każdego oszacowania współczynnika.
- **Przedział ufności.** Wyświetla przedział ufności dla każdego oszacowania współczynnika. Poziom ufności dla przedziału jest ustawiany w oknie dialogowym Opcje.
- **Test  $t$ .** Wyświetla test  $t$  każdego oszacowania współczynnika. Hipoteza zerowa dla każdego testu jest taka, że wartość współczynnika wynosi 0.
- **Kowariancje oszacowań parametrów.** Wyświetla oszacowanie macierzy kowariancji dla współczynników modelu.
- **Korelacje oszacowań parametrów.** Wyświetla oszacowanie macierzy korelacji dla współczynników modelu.

- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.

**Dopasowanie modelu.** Wyświetla statystyki  $R^2$  i błąd średniokwadratowy.

**Średnie w populacji dla zmiennej zależnej i współzmiennych.** Wyświetla informacje podsumowujące dotyczące zmiennej zależnej, współzmiennych i czynników.

**Informacja o planie losowania.** Wyświetla informacje podsumowujące dotyczące próby, w tym nieważoną liczebność i rozmiar populacji.

## Próby złożone: Testy hipotez

**Statystyki testu.** Ta grupa umożliwia wybranie typu statystyki używanej do testowania hipotezy. Można wybrać  $F$ , skorygowane  $F$ , chi-kwadrat i skorygowany chi-kwadrat.

**Stopnie swobody w losowaniu.** Ta grupa zapewnia kontrolę nad liczbą stopni swobody projektu losowania, używanych przy obliczaniu wartości  $p$  dla wszystkich statystyk testujących. Wartość bazująca na projekcie losowania to różnica między liczbą jednostek podstawowego losowania a liczbą warstw na pierwszym etapie losowania. Zamiast niej można podać własną liczbę stopni swobody (dodatnią liczbę całkowitą).

**Korekta dla porównań wielokrotnych.** Podczas testowania hipotez z użyciem kontrastów wielokrotnych całościowy poziom istotności można skorygować na podstawie poziomów istotności dla uwzględnionych kontrastów. W tej grupie możliwy jest wybór metody korekty.

- **Najmniejsza istotna różnica.** Ta metoda nie kontroluje ogólnego prawdopodobieństwa odrzucenia hipotez, które stwierdzają, że niektóre kontrasty liniowe różnią się od wartości hipotezy zerowej.
- *Sekwencyjna Sidaka.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sekwencyjna Bonferroniego.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sidak.* Metoda ta zapewnia węższe granice niż test Bonferroniego.
- *Bonferroni.* Metoda ta dostosowuje obserwowany poziom istotności do faktu przeprowadzania wielu porównań.

## Ogólny model liniowy dla prób złożonych: Estymowane średnie

Okno dialogowe Estymowane średnie umożliwia wyświetlenie estymowanych przez model średnich brzegowych dla poziomów czynników i interakcji czynników określonych w podrzędnym oknie dialogowym Model. Można również zażądać, aby wyświetlona została średnia ogólna populacji.

**Składnik.** Szacowane średnie są obliczane dla wybranych czynników i interakcji czynników.

**Kontrast.** Kontrast określa konfigurację testów hipotezy w celu porównania oszacowanych średnich.

- *Proste.* Wybranie tego typu kontrastu powoduje porównanie średniej każdego poziomu ze średnią wybranego poziomu. Ten typ kontrastu jest przydatny szczególnie w przypadku korzystania z grupy kontrolnej.
- *Odchylenie.* Wybranie tego typu kontrastu powoduje porównanie średniej każdego poziomu (prócz kategorii odniesienia) ze średnią wszystkich poziomów (średnią ogólną). Poziomy czynnika mogą mieć dowolną kolejność.
- *Różnica.* Wybranie tego typu kontrastu powoduje porównanie średniej każdego poziomu (prócz pierwszego) ze średnią poprzednich poziomów Typ ten określany jest także mianem odwrotnego kontrastu Helmerta.



- *Helmerta*. Wybranie tego typu kontrastu powoduje porównanie średniej każdego poziomu czynnika (prócz ostatniego) ze średnią poziomów następných.
- *Powtórzonej*. Wybranie tego typu kontrastu powoduje porównanie średniej każdego poziomu (prócz ostatniego) ze średnią poziomów następných.
- *Wielomianowy*. Wybranie tego typu kontrastu powoduje porównanie efektu liniowego, efektu kwadratowego, sześciennego itd. Dla wszystkich kategorii efekt liniowy zawarty jest w pierwszym stopniu swobody; efekt kwadratowy - w drugim stopniu swobody itd. Tego typu kontrasty używane są często do szacowania trendów wielomianowych.

**Kategoria odniesienia.** Kontrast prosty i kontrast odchylenia wymaga kategorii odniesienia lub poziomu czynnika, z którym porównywane będą inne składniki.

## Ogólny model liniowy dla prób złożonych: Zapisz

**Zapisz zmienne.** Ta grupa umożliwia zapisanie przewidzianých przez model wartości i reszty jako nowych zmienných w aktywnym zbiorze danych.

**Eksportuj model jako dane IBM SPSS Statistics.** Zapisuje zbiór danych w formacie IBM SPSS Statistics zawierającym korelację parametru lub macierz kowariancji z oszacowaniami parametrów, błędami standardowymi, wartościami istotności oraz stopniami swobody. Kolejność zmienných w pliku macierzowym jest następująca.

- **rowtype\_.** przyjmuje wartości (i etykiety wartości), COV (kowariancje), CORR (korelacje), EST (oszacowania parametrów), SE (błędy standardowe), SIG (poziomy istotności) i DF (stopnie swobody projektowania próbkowania). Dla każdego parametru modelu istnieje osobna obserwacja z typem wiersza COV (lub CORR), a także osobna obserwacja dla każdego z pozostałych typów wierszy.
- **varname\_.** Przyjmuje wartości P1, P2, ... odpowiadające uporządkowanej liście wszystkich parametrów modelu dla typów wierszy COV lub CORR, z etykietami wartości odpowiadającymi łańcuchom parametrów przedstawionym w tabeli oszacowania parametrów. Te komórki są puste w przypadku wierszy innych typów.
- **P1, P2, ...** Te zmienne odpowiadają uporządkowanej liście wszystkich parametrów modelu z wartościami etykiet odpowiadającymi ciągom znaków parametrów, które są pokazane w tabeli oszacowań parametrów, a ponadto przyjmują wartości zgodnie z typem wiersza. W odniesieniu do parametrów nadmiarowych wszystkie kowariancje są ustawione na zero, korelacje są ustawione na systemowy brak danych; wszystkie oszacowania parametrów są ustawione na zero, a wszystkie błędy standardowe, poziomy istotności i stopnie swobody reszt są ustawione na systemowy brak danych.

*Uwaga:* Tego pliku nie można użyć bezpośrednio do dalszych analiz w innych procedurach, które odczytują plik macierzowy, chyba że te procedury akceptują wszystkie typy wierszy eksportowane w tym przypadku.

**Eksportuj model jako XML.** Jeśli ta opcja jest wybrana, oszacowania parametrów oraz macierz kowariancji parametrów są zapisywane w formacie XML (PMML). Możesz użyć tego pliku modelu do stosowania informacji o modelu do innych plików danych w celach statystycznych. .

## Ogólny model liniowy dla prób złożonych: Opcje

**Braki danych zdefiniowane przez użytkownika.** Wszystkie zmienne planu, a także zmienna zależna oraz wszelkie współzmiennne, muszą zawierać ważne dane. Obserwacje z nieważnymi danymi w dowolnej z tych zmienných planu są wykluczane z analizy. Te elementy pozwalają zdecydować, czy wartości braków danych zdefiniowanych przez użytkownika są traktowane jako prawidłowe wśród warstw, skupień, podpopulacji i zmienných czynnikowych.

**Oszacowanie przedziału ufności.** Jest to poziom przedziału ufności dla oszacowań współczynników i szacowanych średnich brzegowych. Podaj wartość większą lub równą 50 i mniejszą od 100.

## Dodatkowe właściwości komendy CSGLM

Język składni komend umożliwia również:

- Określanie niestandardowych testów efektów w odniesieniu do liniowej kombinacji efektów lub wartości (za pomocą opcji komendy CUSTOM).
- Ustalanie współzmiennych na wartościach innych niż ich średnie podczas obliczania szacowanych średnich brzegowych (za pomocą opcji komendy EMMEANS).
- Określanie metryk kontrastów wielomianowych (za pomocą opcji komendy EMMEANS).
- Określanie wartości tolerancji dla kontroli osobliwości (za pomocą opcji komendy CRITERIA).
- Utworzenie nazw określonych przez użytkownika dla zapisanych zmiennych (za pomocą opcji komendy SAVE).
- Utworzenie tabeli ogólnych funkcji estymowalnych (za pomocą opcji komendy PRINT).

Pełne informacje na temat składni znajdują się w podręczniku *Command Syntax Reference*.

## Regresja logistyczna dla prób złożonych

Procedura regresji logistycznej dla prób złożonych służy do wykonania analizy regresji logistycznej na binarnej lub wielowymiarowej zmiennej zależnej dla prób wybieranych przy użyciu złożonych metod losowania. Opcjonalnie można zlecić analizy podpopulacji.

**Przykład.** Specjalista ds. pożyczek zgromadził rekordy z przeszłości dotyczące klientów, którym udzielono pożyczki w kilku różnych oddziałach, zgodnie ze złożonym projektem. Poprzez wdrożenie planu losowania specjalista chce się dowiedzieć, czy prawdopodobieństwo tego, który klient nie wywiązuje się ze zobowiązań, jest powiązane z wiekiem, historią zatrudnienia i wielkością zadłużenia.

**Statystyki.** Procedura tworzy oszacowania, oszacowania wykładnicze, błędy standardowe, przedziały ufności, testy  $t$ , efekty planu i pierwiastki kwadratowe efektów planu dla parametrów modelu, jak również korelacje i kowariancje między oszacowaniami parametrów. Dostępne są również statystyki Pseudo  $R^2$ , tabele klasyfikacji oraz statystyki opisowe dla zmiennych zależnych i niezależnych.

Wymagania dotyczące danych w regresji logistycznej dla prób złożonych

**Dane.** Zmienna zależna jest zmienną jakościową. Czynniki są typu jakościowego. Współzmiennie są zmiennymi ilościowymi powiązanymi ze zmienną zależną. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje w pliku danych reprezentują próbkę z projektu złożonego, która ma być analizowana zgodnie z ustawieniami w pliku wybranym w oknie dialogowym Plan prób złożonych.

Wykonywanie regresji logistycznej dla prób złożonych

1. Z menu wybierz:

**Analiza > Próby złożone > Regresja logistyczna...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Regresja logistyczna dla prób złożonych wybierz zmienną zależną.

Opcjonalnie można wykonać następujące czynności:

- Wybrać zmienne dla czynników i współzmiennych odpowiednio do danych.
- Określić zmienną, która ma definiować podpopulację. Analiza jest wykonywana tylko dla wybranej kategorii zmiennej podpopulacji.

## Regresja logistyczna dla prób złożonych: Kategoria odniesienia

Domyślnie procedura Regresja logistyczna dla prób złożonych wybiera kategorię o najwyższej wartości jako kategorię odniesienia. To okno dialogowe umożliwia wybranie jako kategorii odniesienia kategorii o najwyższej wartości, o najniższej wartości lub dowolnej wybranej przez użytkownika.

## Model regresji logistycznej dla prób złożonych

**Określenie efektów modelu.** Domyślnie procedura buduje model efektów głównych na podstawie czynników i współzmiennych określonych w głównym oknie dialogowym. Można też zbudować model niestandardowy zawierający efekty interakcji i składniki zagnieżdżone.

Składniki niezagnieżdżone

Dla wybranych czynników i współzmiennych:

**Interakcja.** Dla wszystkich wybranych zmiennych tworzy składnik interakcji najwyższego rzędu.

**Efekty główne.** Dla każdej wybranej zmiennej tworzy składnik efektów głównych.

**Wszystkie 2 rzędu.** Tworzy wszystkie możliwe interakcje drugiego rzędu wybranych zmiennych.

**Wszystkie 3 rzędu.** Tworzy wszystkie możliwe interakcje trzeciego rzędu wybranych zmiennych.

**Wszystkie 4 rzędu.** Tworzy wszystkie możliwe interakcje czwartego rzędu wybranych zmiennych.

**Wszystkie 5 rzędu.** Tworzy wszystkie możliwe interakcje piątego rzędu wybranych zmiennych.

Składniki zagnieżdżone

W tej procedurze można zbudować składniki zagnieżdżone dla modelu. Składniki zagnieżdżone są przydatne do modelowania efektu czynników lub współzmiennych, których wartości nie wchodzi w interakcje z poziomami innego czynnika. Na przykład sieć sklepów spożywczych może analizować zwyczaje zakupowe swoich klientów w kilku sklepach. Ponieważ każdy klient bywa regularnie tylko w jednym z tych sklepów, efekt *Klient* jest **zagnieżdżony w** efekcie *Lokalizacja sklepu*.

Ponadto można uwzględnić efekty interakcji, takie jak składniki wielomianowe z tą samą współzmienną, lub dodać wiele poziomów zagnieżdżenia do składnika zagnieżdżonego.

**Ograniczenia.** W odniesieniu do składników zagnieżdżonych obowiązują następujące ograniczenia:

- Wszystkie czynniki w interakcji muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A^*A$ .
- Wszystkie czynniki w efekcie zagnieżdżonym muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A(A)$ .
- Efekt nie może być zagnieżdżony w obrębie współzmiennnej. A zatem, jeśli  $A$  jest czynnikiem, a  $X$  jest współzmienną, to określenie  $A(X)$  jest niedozwolone.

**Wyraz wolny.** Wyraz wolny jest zwykle uwzględniany w modelu. Jeśli można założyć, że dane przechodzą przez początek układu współrzędnych, wyraz wolny można wyłączyć z modelu. Nawet jeśli model zawiera wyraz wolny, można pominać związane z nim statystyki.

## Regresja logistyczna dla prób złożonych: Statystyki

**Dopasowanie modelu.** Steruje wyświetlaniem statystyk, które mierzą ogólną wydajność modelu.

- **Pseudo R-kwadrat.** Statystyka  $R^2$  z regresji liniowej nie ma dokładnego odpowiednika wśród modeli regresji logistycznej. Zamiast tego, istnieje wiele miar, które próbują naśladować właściwości statystyk  $R^2$ .
- **Tabela klasyfikacji.** Wyświetla tabelaryczne klasyfikacje krzyżowe obserwowanej kategorii według kategorii prognozowanych przez model na zmiennych zależnych.

**Parametry.** Ta grupa umożliwia sterowanie wyświetlaniem statystyk związanych z parametrami modelu.

- **Oszacowanie.** Wyświetla oszacowania współczynników.
- **Oszacowanie wykładnicze.** Wyświetla podstawę logarytmu naturalnego podniesioną do potęgi oszacowań współczynników. Jeśli oszacowanie ma dobre właściwości dla testów statystycznych, oszacowanie wykładnicze, czyli  $\exp(B)$ , jest łatwiejsze do interpretacji.
- **Błąd standardowy.** Wyświetla błąd standardowy dla każdego oszacowania współczynnika.

- **Przedział ufności.** Wyświetla przedział ufności dla każdego oszacowania współczynnika. Poziom ufności dla przedziału jest ustawiany w oknie dialogowym Opcje.
- **Test t.** Wyświetla test  $t$  każdego oszacowania współczynnika. Hipoteza zerowa dla każdego testu jest taka, że wartość współczynnika wynosi 0.
- **Kowariancje oszacowań parametrów.** Wyświetla oszacowanie macierzy kowariancji dla współczynników modelu.
- **Korelacje oszacowań parametrów.** Wyświetla oszacowanie macierzy korelacji dla współczynników modelu.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.

**Statystyki podsumowujące zmienne modelu.** Wyświetla informacje podsumowujące dotyczące zmiennej zależnej, współzmiennych i czynników.

**Informacja o planie losowania.** Wyświetla informacje podsumowujące dotyczące próby, w tym nieważoną liczebność i rozmiar populacji.

## Próby złożone: Testy hipotez

**Statystyki testu.** Ta grupa umożliwia wybranie typu statystyki używanej do testowania hipotezy. Można wybrać  $F$ , skorygowane  $F$ , chi-kwadrat i skorygowany chi-kwadrat.

**Stopnie swobody w losowaniu.** Ta grupa zapewnia kontrolę nad liczbą stopni swobody projektu losowania, używanych przy obliczaniu wartości  $p$  dla wszystkich statystyk testujących. Wartość bazująca na projekcie losowania to różnica między liczbą jednostek podstawowego losowania a liczbą warstw na pierwszym etapie losowania. Zamiast niej można podać własną liczbę stopni swobody (dodatnią liczbę całkowitą).

**Korekta dla porównań wielokrotnych.** Podczas testowania hipotez z użyciem kontrastów wielokrotnych całościowy poziom istotności można skorygować na podstawie poziomów istotności dla uwzględnionych kontrastów. W tej grupie możliwy jest wybór metody korekty.

- **Najmniejsza istotna różnica.** Ta metoda nie kontroluje ogólnego prawdopodobieństwa odrzucenia hipotez, które stwierdzają, że niektóre kontrasty liniowe różnią się od wartości hipotezy zerowej.
- *Sekwencyjna Sidaka.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sekwencyjna Bonferroniego.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sidak.* Metoda ta zapewnia węższe granice niż test Bonferroniego.
- *Bonferroni.* Metoda ta dostosowuje obserwowany poziom istotności do faktu przeprowadzania wielu porównań.

## Regresja logistyczna dla prób złożonych: Ilorazy szans

Okno dialogowe Ilorazy szans umożliwia wyświetlenie oszacowanych przez model ilorazów szans dla określonych czynników i współzmiennych. Dla każdej kategorii zmiennej zależnej, z wyjątkiem kategorii odniesienia, obliczany jest odrębny zbiór ilorazów szans.

**Czynniki.** Dla każdego wybranego czynnika przedstawia stosunek ilorazów szans przy każdej kategorii czynnika do ilorazów szans przy określonej kategorii odniesienia.

**Współzmiennie.** Dla każdej wybranej współzmiennnej wyświetla stosunek ilorazów szans przy średniej wartości współzmiennnej plus określona liczba jednostek zmiany do ilorazu szans przy średniej.

Podczas obliczania ilorazu szans dla czynnika lub współzmiennej procedura przyjmuje stałe wartości pozostałych czynników, równe ich najwyższym poziomom, oraz stałe wartości pozostałych współzmiennych, równe ich średnim. Jeśli czynnik lub współzmienna wchodzi w interakcje z innymi predyktorami w modelu, to ilorazy szans zależą nie tylko od zmiany określonej zmiennej, ale także od wartości zmiennych, z którymi mają miejsce te interakcje. Jeśli określona współzmienna w modelu wchodzi w interakcje sama ze sobą (na przykład *wiek\*wiek*), to ilorazy szans zależą zarówno od zmiany współzmiennej, jak i od jej wartości.

## Regresja logistyczna dla prób złożonych: Zapisz

**Zapisz zmienne.** Ta grupa umożliwia zapisanie kategorii przewidzianej przez model i przewidzianych prawdopodobieństw jako nowych zmiennych w aktywnym zbiorze danych.

**Eksportuj model jako dane IBM SPSS Statistics.** Zapisuje zbiór danych w formacie IBM SPSS Statistics zawierającym korelację parametru lub macierz kowariancji z oszacowaniami parametrów, błędami standardowymi, wartościami istotności oraz stopniami swobody. Kolejność zmiennych w pliku macierzowym jest następująca.

- **rowtype\_.** przyjmuje wartości (i etykiety wartości), COV (kowariancje), CORR (korelacje), EST (oszacowania parametrów), SE (błędy standardowe), SIG (poziomy istotności) i DF (stopnie swobody projektowania próbkowania). Dla każdego parametru modelu istnieje osobna obserwacja z typem wiersza COV (lub CORR), a także osobna obserwacja dla każdego z pozostałych typów wierszy.
- **varname\_.** Przyjmuje wartości P1, P2, ... odpowiadające uporządkowanej liście wszystkich parametrów modelu dla typów wierszy COV lub CORR, z etykietami wartości odpowiadającymi łańcuchom parametrów przedstawionym w tabeli oszacowania parametrów. Te komórki są puste w przypadku wierszy innych typów.
- **P1, P2, ...** Te zmienne odpowiadają uporządkowanej liście wszystkich parametrów modelu z wartościami etykiet odpowiadającymi ciągom znaków parametrów, które są pokazane w tabeli oszacowań parametrów, a ponadto przyjmują wartości zgodnie z typem wiersza. W odniesieniu do parametrów nadmiarowych wszystkie kowariancje są ustawione na zero, korelacje są ustawione na systemowy brak danych; wszystkie oszacowania parametrów są ustawione na zero, a wszystkie błędy standardowe, poziomy istotności i stopnie swobody reszt są ustawione na systemowy brak danych.

*Uwaga:* Tego pliku nie można użyć bezpośrednio do dalszych analiz w innych procedurach, które odczytują plik macierzowy, chyba że te procedury akceptują wszystkie typy wierszy eksportowane w tym przypadku.

**Eksportuj model jako XML.** Jeśli ta opcja jest wybrana, oszacowania parametrów oraz macierz kowariancji parametrów są zapisywane w formacie XML (PMML). Możesz użyć tego pliku modelu do stosowania informacji o modelu do innych plików danych w celach statystycznych. .

## Regresja logistyczna prób złożonych: Opcje

**Oszacowanie.** Ta grupa umożliwia sterowanie różnymi kryteriami używanymi w estymacji modelu.

- **Maksymalna liczba powtórzeń.** Maksymalna liczba iteracji, jakie wykona algorytm. Podaj nieujemną liczbę całkowitą.
- **Maksimum kroków połowienia.** Przy każdej iteracji rozmiar kroku jest zmniejszany o 0,5, aż do momentu wzrostu logarytmu wiarygodności lub osiągnięcia maksymalnej wartości połowienia kroków. Podaj dodatnią liczbę całkowitą.
- **Ogranicz iteracje ze względu na zmiany w ocenach parametrów.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana oszacowania parametru jest mniejsza od określonej wartości, która musi być nieujemna.
- **Ogranicz iteracje z powodu zmian w logarytmie wiarygodności.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana funkcji logarytmu wiarygodności jest mniejsza od określonej wartości, która musi być nieujemna.
- **Sprawdzaj rozdział punktów danych.** Po wybraniu tej opcji algorytm przeprowadza testy, aby sprawdzić, czy oszacowania parametru mają unikalne wartości. Rozdzielenie następuje, kiedy procedura może stworzyć model, który poprawnie klasyfikuje wszystkie obserwacje.

- **Pokaż przebieg iteracji zgodnie z przyrostem kroków.** Wyświetla oszacowania parametrów i statystyki w przypadku wszystkich iteracji  $n$  rozpoczynających się od iteracji  $0^{sz}$  (szacunki początkowe). Jeśli zostanie wybrana opcja pokazywania przebiegu iteracji, wówczas zawsze przedstawiana będzie ostatnia iteracja bez względu na wartość  $n$ .

**Braki danych zdefiniowane przez użytkownika.** Wszystkie zmienne planu, a także zmienna zależna oraz wszelkie współzmiennie, muszą zawierać ważne dane. Obserwacje z nieważnymi danymi w dowolnej z tych zmiennych planu są wykluczane z analizy. Te elementy pozwalają zdecydować, czy wartości braków danych zdefiniowanych przez użytkownika są traktowane jako prawidłowe wśród warstw, skupień, podpopulacji i zmiennych czynnikowych.

**Oszacowanie przedziału ufności.** Jest to poziom przedziału ufności dla oszacowań współczynników, wykładniczego oszacowania współczynników i ilorazów szans. Podaj wartość większą lub równą 50 i mniejszą od 100.

## Dodatkowe właściwości komendy CSLOGISTIC

Język składni komend umożliwia również:

- Określanie niestandardowych testów efektów w odniesieniu do liniowej kombinacji efektów lub wartości (za pomocą opcji komendy CUSTOM).
- Poprawianie wartości innych zmiennych modelu podczas obliczania ilorazów szans dla czynników i współzmiennych (za pomocą opcji komendy ODDS RATIOS).
- Określanie wartości tolerancji dla kontroli osobliwości (za pomocą opcji komendy CRITERIA).
- Utworzenie nazw określonych przez użytkownika dla zapisanych zmiennych (za pomocą opcji komendy SAVE).
- Utworzenie tabeli ogólnych funkcji estymowalnych (za pomocą opcji komendy PRINT).

Pełne informacje na temat składni znajdują się w podręczniku *Command Syntax Reference*.

## Regresja porządkowa prób złożonych

Procedura regresji porządkowej prób złożonych służy do analizy regresji binarnej lub porządkowej zmiennej zależnej dla próbek pobieranych przy użyciu złożonych metod doboru próby. Opcjonalnie można zlecić analizy podpopulacji.

**Przykład.** Przedstawiciele rozpatrujący projekt ustawy przed przekazaniem go do ustawodawcy zastanawiają się, czy istnieje publiczne poparcie dla tego projektu oraz w jaki sposób poparcie dla projektu jest powiązane z danymi demograficznymi głosujących. Ankieeterzy tworzą i przeprowadzają wywiady, zgodnie ze złożonym planem logowania. Zastosowanie regresji porządkowej prób złożonych umożliwia dopasowanie modelu do poziomu poparcia dla projektu ustawy na podstawie danych demograficznych głosujących.

Wymagania dotyczące danych dla regresji porządkowej prób złożonych

**Dane.** Zmienna zależna jest porządkowa. Czynniki są typu jakościowego. Współzmiennie są zmiennymi ilościowymi powiązane ze zmienną zależną. Zmienne podpopulacji mogą być łańcuchowe lub numeryczne, ale powinny być jakościowe.

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym [Plan prób złożonych](#).

Wykonywanie regresji porządkowej prób złożonych

1. Z menu wybierz:

**Analiza > Próby złożone > Regresja porządkowa...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.
3. Kliknij przycisk **Kontynuuj**.

4. W oknie dialogowym Regresja porządkowa prób złożonych wybierz zmienną zależną.

Opcjonalnie można wykonać następujące czynności:

- Wybrać zmienne dla czynników i współzmiennych odpowiednio do danych.
- Określić zmienną, która ma definiować podpopulację. Analiza jest wykonywana tylko dla wybranej kategorii zmiennej podpopulacji, ale wariancje są wciąż poprawnie szacowane na podstawie całego zbioru danych.
- Wybierz funkcję łączenia.

**Funkcja łączenia.** Funkcja łączenia to transformacja skumulowanych prawdopodobieństw, która umożliwia estymację modelu. Dostępnych jest pięć funkcji łączenia.

- **Logit.**  $f(x)=\log(x/(1-x))$ . Zwykle używane w przypadku kategorii równomiernie rozproszonych.
- **Komplementarny logarytmiczno-logarytmiczny.**  $f(x)=\log(-\log(1-x))$ . Zazwyczaj używane, gdy wyższe kategorie są bardziej prawdopodobne.
- **Ujemny logarytmiczno-logarytmiczny.**  $f(x)=-\log(-\log(x))$ . Zazwyczaj używane, gdy niższe kategorie są bardziej prawdopodobne.
- **Probit.**  $f(x)=\Phi^{-1}(x)$ . Zazwyczaj używane, gdy zmienna utajona ma rozkład normalny.
- **Cauchit (odwrócony Cauchy'ego).**  $f(x)=\tan(\pi(x-0,5))$ . Zazwyczaj używane, gdy zmienna utajona zawiera wiele wartości skrajnych.

## Regresja porządkowa prób złożonych: Prawdopodobieństwa odpowiedzi

Okno dialogowe Prawdopodobieństwa odpowiedzi umożliwia określenie, czy skumulowane prawdopodobieństwo odpowiedzi (tj. prawdopodobieństwo przynależności do konkretnej kategorii zmiennej zależnej i wszystkich poprzednich kategorii) zwiększa się wraz ze wzrostem lub spadkiem wartości zmiennej zależnej.

## Regresja porządkowa prób złożonych: Model

**Określenie efektów modelu.** Domyślnie procedura buduje model efektów głównych na podstawie czynników i współzmiennych określonych w głównym oknie dialogowym. Można też zbudować model niestandardowy zawierający efekty interakcji i składniki zagnieżdżone.

Składniki niezagnieżdżone

Dla wybranych czynników i współzmiennych:

**Interakcja.** Dla wszystkich wybranych zmiennych tworzy składnik interakcji najwyższego rzędu.

**Efekty główne.** Dla każdej wybranej zmiennej tworzy składnik efektów głównych.

**Wszystkie 2 rzędu.** Tworzy wszystkie możliwe interakcje drugiego rzędu wybranych zmiennych.

**Wszystkie 3 rzędu.** Tworzy wszystkie możliwe interakcje trzeciego rzędu wybranych zmiennych.

**Wszystkie 4 rzędu.** Tworzy wszystkie możliwe interakcje czwartego rzędu wybranych zmiennych.

**Wszystkie 5 rzędu.** Tworzy wszystkie możliwe interakcje piątego rzędu wybranych zmiennych.

Składniki zagnieżdżone

W tej procedurze można zbudować składniki zagnieżdżone dla modelu. Składniki zagnieżdżone są przydatne do modelowania efektu czynników lub współzmiennych, których wartości nie wchodzi w interakcje z poziomami innego czynnika. Na przykład sieć sklepów spożywczych może analizować zwyczaje zakupowe swoich klientów w kilku sklepach. Ponieważ każdy klient bywa regularnie tylko w jednym z tych sklepów, efekt *Klient* jest **zagnieżdżony w** efekcie *Lokalizacja sklepu*.

Ponadto można uwzględnić efekty interakcji, takie jak składniki wielomianowe z tą samą współzmienną, lub dodać wiele poziomów zagnieżdżenia do składnika zagnieżdżonego.

**Ograniczenia.** W odniesieniu do składników zagnieżdżonych obowiązują następujące ograniczenia:

- Wszystkie czynniki w interakcji muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A^*A$ .
- Wszystkie czynniki w efekcie zagnieżdżonym muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A(A)$ .
- Efekt nie może być zagnieżdżony w obrębie współzmiennnej. A zatem, jeśli  $A$  jest czynnikiem, a  $X$  jest współzmienną, to określenie  $A(X)$  jest niedozwolone.

## Regresja porządkowa prób złożonych: Statystyki

**Dopasowanie modelu.** Steruje wyświetlaniem statystyk, które mierzą ogólną wydajność modelu.

- **Pseudo R-kwadrat.** Statystyka  $R^2$  z regresji liniowej nie ma dokładnego odpowiednika wśród modeli regresji porządkowej. Zamiast tego, istnieje wiele miar, które próbują naśladować właściwości statystyk  $R^2$ .
- **Tabela klasyfikacji.** Wyświetla tabelaryczne klasyfikacje krzyżowe obserwowanej kategorii według kategorii prognozowanych przez model na zmiennych zależnych.

**Parametry.** Ta grupa umożliwia sterowanie wyświetlaniem statystyk związanych z parametrami modelu.

- **Oszacowanie.** Wyświetla oszacowania współczynników.
- **Oszacowanie wykładnicze.** Wyświetla podstawę logarytmu naturalnego podniesioną do potęgi oszacowań współczynników. Jeśli oszacowanie ma dobre właściwości dla testów statystycznych, oszacowanie wykładnicze, czyli  $\exp(B)$ , jest łatwiejsze do interpretacji.
- **Błąd standardowy.** Wyświetla błąd standardowy dla każdego oszacowania współczynnika.
- **Przedział ufności.** Wyświetla przedział ufności dla każdego oszacowania współczynnika. Poziom ufności dla przedziału jest ustawiany w oknie dialogowym Opcje.
- **Test t.** Wyświetla test  $t$  każdego oszacowania współczynnika. Hipoteza zerowa dla każdego testu jest taka, że wartość współczynnika wynosi 0.
- **Kowariancje oszacowań parametrów.** Wyświetla oszacowanie macierzy kowariancji dla współczynników modelu.
- **Korelacje oszacowań parametrów.** Wyświetla oszacowanie macierzy korelacji dla współczynników modelu.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara, wyrażona w jednostkach porównywalnych z jednostkami błędów standardowych, efektu określenia złożonego planu, przy czym wartości dalsze od 1 wskazują na większy efekt.

**Linie równoległe.** W tej grupie można zażądać statystyk związanych z modelem o nierównoległych liniach, w którym dla każdej kategorii odpowiedzi (z wyjątkiem ostatniej) jest dopasowana oddzielna linia regresji.

- **Test Walda.** Umożliwia przetestowanie hipotezy zerowej twierzącej, że parametry regresji są równe dla wszystkich skumulowanych odpowiedzi. Model z liniami nierównoległymi jest szacowany i stosowany jest test Walda o równych parametrach.
- **Oceny parametrów.** Wyświetla oszacowania współczynników i błędy standardowe dla modelu z liniami nierównoległymi.
- **Kowariancje oszacowań parametrów.** Wyświetla oszacowanie macierzy kowariancji dla współczynników modelu z liniami nierównoległymi.

**Statystyki podsumowujące zmienne modelu.** Wyświetla informacje podsumowujące dotyczące zmiennej zależnej, współzmiennych i czynników.

**Informacja o planie losowania.** Wyświetla informacje podsumowujące dotyczące próby, w tym nieważoną liczebność i rozmiar populacji.



## Próby złożone: Testy hipotez

**Statystyki testu.** Ta grupa umożliwia wybranie typu statystyki używanej do testowania hipotezy. Można wybrać  $F$ , skorygowane  $F$ , chi-kwadrat i skorygowany chi-kwadrat.

**Stopnie swobody w losowaniu.** Ta grupa zapewnia kontrolę nad liczbą stopni swobody projektu losowania, używanych przy obliczaniu wartości  $p$  dla wszystkich statystyk testujących. Wartość bazująca na projekcie losowania to różnica między liczbą jednostek podstawowego losowania a liczbą warstw na pierwszym etapie losowania. Zamiast niej można podać własną liczbę stopni swobody (dodatnią liczbę całkowitą).

**Korekta dla porównań wielokrotnych.** Podczas testowania hipotez z użyciem kontrastów wielokrotnych całościowy poziom istotności można skorygować na podstawie poziomów istotności dla uwzględnionych kontrastów. W tej grupie możliwy jest wybór metody korekty.

- **Najmniejsza istotna różnica.** Ta metoda nie kontroluje ogólnego prawdopodobieństwa odrzucenia hipotez, które stwierdzają, że niektóre kontrasty liniowe różnią się od wartości hipotezy zerowej.
- *Sekwencyjna Sidaka.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sekwencyjna Bonferroniego.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sidak.* Metoda ta zapewnia węższe granice niż test Bonferroniego.
- *Bonferroni.* Metoda ta dostosowuje obserwowany poziom istotności do faktu przeprowadzania wielu porównań.

## Regresja porządkowa prób złożonych: Ilorazy szans

Okno dialogowe Ilorazy szans umożliwia wyświetlenie oszacowanych przez model skumulowanych ilorazów szans dla określonych czynników i współzmiennych. Ta funkcja jest dostępna tylko w przypadku modeli korzystających z funkcji łączenia logit. Jeden skumulowany iloraz szans jest obliczany dla wszystkich kategorii zmiennej zależnej z wyjątkiem ostatniej kategorii; w modelu proporcjonalnym ilorazów szans postuluje się, że wszystkie one są równe.

**Czynniki.** Dla każdego wybranego czynnika przedstawia stosunek skumulowanych ilorazów szans przy każdej kategorii czynnika do ilorazów szans przy określonej kategorii odniesienia.

**Współzmiennie.** Dla każdej wybranej współzmiennej wyświetla stosunek skumulowanych ilorazów szans przy średniej wartości współzmiennej plus określona liczba jednostek zmiany do ilorazu szans przy średniej.

Podczas obliczania ilorazu szans dla czynnika lub współzmiennej procedura przyjmuje stałe wartości pozostałych czynników, równe ich najwyższym poziomom, oraz stałe wartości pozostałych współzmiennych, równe ich średnim. Jeśli czynnik lub współzmienna wchodzi w interakcje z innymi predyktorami w modelu, to ilorazy szans zależą nie tylko od zmiany określonej zmiennej, ale także od wartości zmiennych, z którymi mają miejsce te interakcje. Jeśli określona współzmienna w modelu wchodzi w interakcje sama ze sobą (na przykład *wiek\*wiek*), to ilorazy szans zależą zarówno od zmiany współzmiennej, jak i od jej wartości.

## Regresja porządkowa prób złożonych: Zapisz

**Zapisz zmienne.** Ta grupa umożliwia zapisanie przewidzianej przez model kategorii, prawdopodobieństwa przewidzianej kategorii, prawdopodobieństwa zaobserwowanej kategorii, prawdopodobieństw skumulowanych oraz przewidywanych prawdopodobieństw jako nowych zmiennych w aktywnym zbiorze danych.

**Eksportuj model jako dane IBM SPSS Statistics.** Zapisuje zbiór danych w formacie IBM SPSS Statistics zawierającym korelację parametru lub macierz kowariancji z oszacowaniami parametrów,

błędami standardowymi, wartościami istotności oraz stopniami swobody. Kolejność zmiennych w pliku macierzowym jest następująca.

- **rowtype\_.** przyjmuje wartości (i etykiety wartości), COV (kowariancje), CORR (korelacje), EST (oszacowania parametrów), SE (błędy standardowe), SIG (poziomy istotności) i DF (stopnie swobody projektowania próbkowania). Dla każdego parametru modelu istnieje osobna obserwacja z typem wiersza COV (lub CORR), a także osobna obserwacja dla każdego z pozostałych typów wierszy.
- **varname\_.** Przyjmuje wartości P1, P2, ... odpowiadające uporządkowanej liście wszystkich parametrów modelu dla typów wierszy COV lub CORR, z etykietami wartości odpowiadającymi łańcuchom parametrów przedstawionym w tabeli oszacowania parametrów. Te komórki są puste w przypadku wierszy innych typów.
- **P1, P2, ...** Te zmienne odpowiadają uporządkowanej liście wszystkich parametrów modelu z wartościami etykiet odpowiadającymi ciągom znaków parametrów, które są pokazane w tabeli oszacowań parametrów, a ponadto przyjmują wartości zgodnie z typem wiersza. W odniesieniu do parametrów nadmiarowych wszystkie kowariancje są ustawione na zero, korelacje są ustawione na systemowy brak danych; wszystkie oszacowania parametrów są ustawione na zero, a wszystkie błędy standardowe, poziomy istotności i stopnie swobody reszt są ustawione na systemowy brak danych.

*Uwaga:* Tego pliku nie można użyć bezpośrednio do dalszych analiz w innych procedurach, które odczytują plik macierzowy, chyba że te procedury akceptują wszystkie typy wierszy eksportowane w tym przypadku.

**Eksportuj model jako XML.** Jeśli ta opcja jest wybrana, oszacowania parametrów oraz macierz kowariancji parametrów są zapisywane w formacie XML (PMML). Możesz użyć tego pliku modelu do stosowania informacji o modelu do innych plików danych w celach statystycznych. .

## Regresja porządkowa prób złożonych: Opcje

**Metoda oszacowania.** Można wybrać metodę oszacowania parametru: Newtona-Raphsona, Fishera lub metodę hybrydową, w której interakcje oceny Fishera są wykonywane przed przejściem do metody Newtona-Raphsona. Jeśli zbieżność zostanie osiągnięta w fazie oceny Fishera metodą hybrydową przed wykonaniem maksymalnej liczby iteracji Fishera, algorytm będzie kontynuował działanie, stosując metodę Newtona-Raphsona.

**Oszacowanie.** Ta grupa umożliwia sterowanie różnymi kryteriami używanymi w estymacji modelu.

- **Maksymalna liczba powtórzeń.** Maksymalna liczba iteracji, jakie wykona algorytm. Podaj nieujemną liczbę całkowitą.
- **Maksimum kroków połowienia.** Przy każdej iteracji rozmiar kroku jest zmniejszany o 0,5, aż do momentu wzrostu logarytmu wiarygodności lub osiągnięcia maksymalnej wartości połowienia kroków. Podaj dodatnią liczbę całkowitą.
- **Ogranicz iteracje ze względu na zmiany w ocenach parametrów.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana oszacowania parametru jest mniejsza od określonej wartości, która musi być nieujemna.
- **Ogranicz iteracje z powodu zmian w logarytmie wiarygodności.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana funkcji logarytmu wiarygodności jest mniejsza od określonej wartości, która musi być nieujemna.
- **Sprawdzaj rozdział punktów danych.** Po wybraniu tej opcji algorytm przeprowadza testy, aby sprawdzić, czy oszacowania parametru mają unikalne wartości. Rozdzielenie następuje, kiedy procedura może stworzyć model, który poprawnie klasyfikuje wszystkie obserwacje.
- **Pokaż przebieg iteracji zgodnie z przyrostem kroków.** Wyświetla oszacowania parametrów i statystyki w przypadku wszystkich iteracji  $n$  rozpoczynających się od iteracji 0 (szacunki początkowe). Jeśli zostanie wybrana opcja pokazywania przebiegu iteracji, wówczas zawsze przedstawiana będzie ostatnia iteracja bez względu na wartość  $n$ .

**Braki danych zdefiniowane przez użytkownika.** Ilościowe zmienne planu, a także zmienna zależna oraz wszelkie współzmiennne, powinny zawierać ważne dane. Obserwacje z nieważnymi danymi w dowolnej z tych zmiennych planu są wykluczane z analizy. Te elementy pozwalają zdecydować, czy wartości

braków danych zdefiniowanych przez użytkownika są traktowane jako prawidłowe wśród warstw, skupień, podpopulacji i zmiennych czynnikowych.

**Oszacowanie przedziału ufności.** Jest to poziom przedziału ufności dla oszacowań współczynników, wykładniczego oszacowania współczynników i ilorazów szans. Podaj wartość większą lub równą 50 i mniejszą od 100.

## Dodatkowe właściwości komendy CSORDINAL

Język składni komend umożliwia również:

- Określanie niestandardowych testów efektów w odniesieniu do liniowej kombinacji efektów lub wartości (za pomocą opcji komendy CUSTOM).
- Ustalenie wartości innych zmiennych modelu na wartości inne niż ich średnie podczas obliczania skumulowanych ilorazów szans dla czynników i współzmiennych (przy użyciu opcji komendy ODDSRATIOS).
- Użycie wartości bez etykiet jako niestandardowych kategorii odniesienia dla czynników, gdy wymagane są ilorazy szans (przy użyciu opcji komendy ODDSRATIOS).
- Określanie wartości tolerancji dla kontroli osobliwości (za pomocą opcji komendy CRITERIA).
- Utworzenie tabeli ogólnych funkcji estymowalnych (za pomocą opcji komendy PRINT).
- Zapisanie więcej niż 25 zmiennych prawdopodobieństwa (przy użyciu opcji komendy SAVE).

Pełne informacje na temat składni znajdują się w podręczniku *Command Syntax Reference*.

## regresja Coxa prób złożonych

Procedura regresji Coxa prób złożonych przeprowadza analizę przeżycia dla próbek pobranych za pomocą złożonych metod pobierania próbek. Opcjonalnie można zlecić analizy podpopulacji.

**Przykłady.** Rządowa agencja ścigania jest zaniepokojona stopami recydywistów w ich obszarze jurysdykcji. Jednym ze środków recydywy jest czas do drugiego aresztowania za sprawców. Agencja chciałaby wzorować czas na aresztowanie za pomocą regresji Coxa, ale obawiają się, że założenie o proporcjonalnym zagrożeniu jest nieważne w kategoriach wiekowych.

Badacze medycy badają czasy przeżycia pacjentów, którzy prowadzą program rehabilitacyjny po udarze niedokrwiennej serca. Istnieje możliwość wystąpienia wielu przypadków na jeden temat, ponieważ historie pacjentów zmieniają się jako występowanie istotnych zdarzeń niezgonów, a czasy tych zdarzeń odnotowane. Próbka jest również lewa-obcięta w tym sensie, że obserwowane czasy przeżycia są "napompowane" przez długość rehabilitacji, ponieważ podczas gdy początek ryzyka zaczyna się w momencie udaru niedokrwionego, tylko pacjenci, którzy przeżyli obok programu rehabilitacji, są w próbie.

Uwagi dotyczące danych regresji Coxa z próbkami złożonymi

**Czas przeżycia.** Procedura stosuje regresję Coxa do analizy czasów przeżycia-czyli długości czasu przed wystąpieniem zdarzenia. Istnieją dwa sposoby określenia czasu przeżycia, w zależności od czasu rozpoczęcia przedziału czasu:

- **Czas=0.** Zwykle użytkownik będzie miał pełne informacje na temat początku przedziału dla każdego tematu i po prostu będzie miał zmienną zawierającą czasy zakończenia (lub utworzyć jedną zmienną z czasem zakończenia ze zmiennych Data i godzina; patrz poniżej).
- **Zależy od tematu.** Jest to właściwe w przypadku **obcinanie po lewej stronie**, nazywanych również **opóźniony wpis**; na przykład, jeśli analizujesz czasy przeżycia pacjentów wychodzenia z programu rehabilitacja po udarze, można uznać, że ich początek ryzyka rozpoczyna się w momencie udaru mózgu. Jeśli jednak Twoja próba obejmuje tylko pacjentów, którzy przeżyli program rehabilitacji, to Twoja próba zostaje obcięta w tym sensie, że obserwowane czasy przeżycia są "napompowane" przez długość rehabilitacji. Można to uwzględnić, określając czas, w którym wpleciono z rehabilitacji jako czas wejścia do badania.

**Zmienne daty i godziny.** Zmienne typu Data i godzina nie mogą być używane do bezpośredniego definiowania początku i końca przedziału czasu. Jeśli istnieją zmienne typu Data i godzina, należy użyć ich do utworzenia zmiennych zawierających czasy przeżycia. Jeśli nie ma lewego obcięcia, po prostu utwórz zmienną zawierającą czasy zakończenia w oparciu o różnicę między datą wprowadzenia do badania a datą obserwacji. Jeśli zostanie zastosowane obcięcie w lewo, utwórz zmienną zawierającą czasy rozpoczęcia w oparciu o różnicę między datą rozpoczęcia badania a datą wprowadzenia, a zmienną zawierającą czasy zakończenia, na podstawie różnicy między datą rozpoczęcia badania a datą obserwacji.

**Status zdarzenia.** Potrzebna jest zmienna, która rejestruje, czy w danym okresie zainteresował się interesującego nas temat. Osoby, dla których zdarzenie nie miało miejsca, mają prawo-ocenzurowane.

**Identyfikator podmiotu.** W prosty sposób można uwzględnić stałe odcinkowe, zależne od czasu, dzieląc obserwacje na jeden temat w wielu przypadkach. Na przykład, jeśli analizujesz czasy przeżycia dla pacjentów po udarze mózgu, zmienne reprezentujące ich historię medyczną powinny być użyteczne jako predyktory. Z czasem mogą doświadczyć poważnych zdarzeń medycznych, które zmieniają swoją historię medyczną. W poniższej tabeli przedstawiono sposób struktury takiego zbioru danych: *ID pacjenta* jest identyfikatorem podmiotu, *Godzina zakończenia* definiuje obserwowane przedziały, *Status* rejestruje ważne zdarzenia medyczne, a *Wcześniejsza historia zawału serca* i *Wcześniejsza historia krwotoku* to predyktory zależne od czasu, stałe odcinkowe.

*Tabela 1. Struktura danych dla uwzględniania predyktorów zależnych od czasu pracy stałej*

<i>ID pacjenta</i>	<i>Godzina zakończenia</i>	<i>Status</i>	<i>Wcześniejsza historia zawału serca</i>	<i>Wcześniejsza historia krwotoku</i>
1	5	Atak serca	Nie	Nie
1	7	Krwotoczne	Tak	Nie
1	8	Zmarł	Tak	Tak
2	24	Zmarł	Nie	Nie
3	8	Atak serca	Nie	Nie
3	15	Zmarł	Tak	Nie

**Założenia.** Obserwacje z pliku danych reprezentują próbę ze złożonego projektu, która powinna zostać poddana analizie zgodnie ze specyfikacjami określonymi w pliku wybranym w oknie dialogowym Plan prób złożonych.

Zazwyczaj modele regresji Coxa zakładają proporcjonalne zagrożenia-to znaczy stosunek zagrożeń z jednego przypadku do drugiego nie powinien się zmieniać w czasie. Jeśli to założenie nie jest wstrzymane, może być konieczne dodanie predyktorów zależnych od czasu do modelu.

**Analiza Kaplana-Meiera.** Jeśli użytkownik nie wybierze żadnych predyktorów (lub nie wprowadzi żadnych wybranych predyktorów do modelu) i wybierze metodę limitu produktu na potrzeby obliczania krzywej przeżycia według planu bazowego na karcie Opcje, procedura przeprowadzi analizę przeżycia typu Kaplana-Meiera.

Aby uzyskać złożoną regresję Coxa,

1. Z menu wybierz:

**Analizuj > Przykłady złożone > Regresja porządkowa ...**

2. Wybierz plik planu. Opcjonalnie można też wybrać plik niestandardowych prawdopodobieństw łącznych.

3. Kliknij przycisk **Kontynuuj**.

4. Określ czas przeżycia, wybierając czas wejścia i wyjścia z badania.

5. Wybierz zmienną statusu zdarzenia.

6. Kliknij opcję Zdefiniuj zdarzenie i zdefiniuj co najmniej jedną wartość zdarzenia.

Opcjonalnie można wybrać identyfikator podmiotu.

## Definiowanie zdarzenia

Określ wartości, które wskazują wystąpienie zdarzenia końcowego.

- **Poszczególne wartości.** Określ jedną lub więcej wartości, wprowadzając je do siatki lub wybierając je z listy wartości o zdefiniowanych etykietach wartości.
- **Zakres wartości.** Określ zakres wartości, wprowadzając wartości minimalne i maksymalne lub wybierając wartości z listy o zdefiniowanych etykietach wartości.

## Predyktory

Karta Predyktory umożliwia określenie czynników i współzmiennych używanych do budowania efektów modelu.

**Czynniki.** Czynniki są predyktorami jakościowymi; mogą być liczbowe lub łańcuchowe.

**Współzmiennie.** Współzmiennie są predyktorami skalarnymi; muszą być liczbowe.

**Predyktory zależne od czasu.** Istnieją pewne sytuacje, w których założenia proporcjonalnego hazardu nie są wstrzymane. Oznacza to, że współczynniki hazardu zmieniają się w czasie. Wartości jednej (lub większej liczby) predyktorów są różne w różnych punktach czasowych. W takich przypadkach konieczne jest określenie predyktorów zależnych od czasu. Więcej informacji można znaleźć w temacie [“Definiowanie predyktora zależnego od czasu”](#) na stronie 33. Predyktory zależne od czasu mogą być wybierane jako czynniki lub współzmiennie.

## Definiowanie predyktora zależnego od czasu

Okno dialogowe Zdefiniuj predyktor zależny od czasu umożliwia utworzenie predyktora, który jest zależny od wbudowanej zmiennej czasu  $T_$ . Tej zmiennej można użyć do zdefiniowania współzmiennych zależnych od czasu na dwa ogólne sposoby:

- Aby oszacować rozszerzony model regresji Coxa, który pozwala na nieproporcjonalne hazardy, można to zrobić, definiując predyktor zależny od czasu jako funkcję zmiennej czasu  $T_$  i danej współzmiennnej. Typowym przykładem może być prosty iloraz zmiennej czasu i predyktora, ale można również określić bardziej złożone funkcje.
- Niektóre zmienne mogą mieć różne wartości w różnych punktach czasowych, nie będąc systematycznie zależne od czasu. W takich przypadkach konieczne jest zdefiniowanie **segmentowanego predyktora zależnego od czasu**, co można zrobić przy użyciu wyrażeń logicznych. Wyrażenie logiczne przyjmuje wartość 1, jeśli jest prawdziwe, albo 0, jeśli jest fałszywe. Korzystając z serii wyrażeń logicznych, można utworzyć predyktor zależny od czasu z zestawu pomiarów. Na przykład, jeśli ciśnienie krwi jest mierzone raz w tygodniu przez cztery tygodnie badania (identyfikowane jako  $BP1$  do  $BP4$ ), można zdefiniować predyktor zależny od czasu jako  $(T_ < 1) * BP1 + (T_ \geq 1 \ \& \ T_ < 2) * BP2 + (T_ \geq 2 \ \& \ T_ < 3) * BP3 + (T_ \geq 3 \ \& \ T_ < 4) * BP4$ . Zauważ, że dla każdego przypadku dokładnie jeden z członów w nawiasach będzie równy 1, a reszta będzie równa 0. Innymi słowy, ta funkcja oznacza, że jeśli czas jest krótszy niż jeden tydzień, należy użyć  $BP1$ ; jeśli jest to więcej niż jeden tydzień, ale mniej niż dwa tygodnie, należy użyć  $BP2$  itd.

*Uwaga:* Jeśli segmentowany predyktor zależny od czasu jest stały w segmentach, tak jak w powyższym przykładzie ciśnienia krwi, można ułatwić określanie stałego predyktora zależnego od czasu, dzieląc obiekty na wiele przypadków. Więcej informacji na ten temat można znaleźć w dyskusji na temat identyfikatorów tematów w [“regresja Coxa prób złożonych”](#) na stronie 31.

W oknie dialogowym definiowania predyktora zależnego od czasu można użyć elementów sterujących budowaniem funkcji w celu zbudowania wyrażenia dla współzmiennnej zależnej od czasu lub można ją wprowadzić bezpośrednio w obszarze tekstowym Wyrażenie liczbowe. Należy zauważyć, że stałe łańcuchowe muszą być ujęte w cudzysłowy lub apostrofy, a stałe liczbowe muszą być wpisane w formacie amerykańskim, z kropką dziesiętną. Wynikowa zmienna ma nazwę określoną przez użytkownika i powinna być uwzględniona jako czynnik lub współzmienna na karcie Predyktory.

## Podgrupy

**Warstwa podstawowa.** Dla każdej wartości tej zmiennej obliczana jest osobna linia bazowa hazardu i funkcja przeżycia, podczas gdy jeden zestaw współczynników modelu jest szacowany na podstawie warstw.

**Zmienna podpopulacji.** Określić zmienną, która ma definiować podpopulację. Analiza jest wykonywana tylko dla wybranej kategorii zmiennej podpopulacji.

## Model

**Określenie efektów modelu.** Domyślnie procedura buduje model efektów głównych na podstawie czynników i współzmiennych określonych w głównym oknie dialogowym. Można też zbudować model niestandardowy zawierający efekty interakcji i składniki zagnieżdżone.

Składniki niezagnieżdżone

Dla wybranych czynników i współzmiennych:

**Interakcja.** Dla wszystkich wybranych zmiennych tworzy składnik interakcji najwyższego rzędu.

**Efekty główne.** Dla każdej wybranej zmiennej tworzy składnik efektów głównych.

**Wszystkie 2 rzędu.** Tworzy wszystkie możliwe interakcje drugiego rzędu wybranych zmiennych.

**Wszystkie 3 rzędu.** Tworzy wszystkie możliwe interakcje trzeciego rzędu wybranych zmiennych.

**Wszystkie 4 rzędu.** Tworzy wszystkie możliwe interakcje czwartego rzędu wybranych zmiennych.

**Wszystkie 5 rzędu.** Tworzy wszystkie możliwe interakcje piątego rzędu wybranych zmiennych.

Składniki zagnieżdżone

W tej procedurze można zbudować składniki zagnieżdżone dla modelu. Składniki zagnieżdżone są przydatne do modelowania efektu czynników lub współzmiennych, których wartości nie wchodzą w interakcje z poziomami innego czynnika. Na przykład sieć sklepów spożywczych może analizować zwyczaje zakupowe swoich klientów w kilku sklepach. Ponieważ każdy klient bywa regularnie tylko w jednym z tych sklepów, efekt *Klient* jest **zagnieżdżony w** efekcie *Lokalizacja sklepu*.

Ponadto można uwzględnić efekty interakcji, takie jak składniki wielomianowe z tą samą współzmienną, lub dodać wiele poziomów zagnieżdżenia do składnika zagnieżdżonego.

**Ograniczenia.** W odniesieniu do składników zagnieżdżonych obowiązują następujące ograniczenia:

- Wszystkie czynniki w interakcji muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A^*A$ .
- Wszystkie czynniki w efekcie zagnieżdżonym muszą być unikalne. A zatem, jeśli  $A$  jest czynnikiem, to niedozwolone jest określenie  $A(A)$ .
- Efekt nie może być zagnieżdżony w obrębie współzmiennnej. A zatem, jeśli  $A$  jest czynnikiem, a  $X$  jest współzmienną, to określenie  $A(X)$  jest niedozwolone.

## Statystyka

**Informacja o planie losowania.** Wyświetla informacje podsumowujące dotyczące próby, w tym nieważoną liczebność i rozmiar populacji.

**Podsumowanie zdarzeń i cenzurowania.** Wyświetla informacje podsumowujące o liczbie i odsetku ocenzurowanych obserwacji.

**Zestaw ryzyka w czasie zdarzenia.** Wyświetla liczbę zdarzeń i liczbę zdarzeń ryzyka dla każdego czasu zdarzenia w każdej warstwie podstawowej.

**Parametry.** Ta grupa umożliwia sterowanie wyświetlaniem statystyk związanych z parametrami modelu.

- **Oszacowanie.** Wyświetla oszacowania współczynników.

- **Oszacowanie wykładnicze.** Wyświetla podstawę logarytmu naturalnego podniesioną do potęgi oszacowań współczynników. Jeśli oszacowanie ma dobre właściwości dla testów statystycznych, oszacowanie wykładnicze, czyli  $\exp(B)$ , jest łatwiejsze do interpretacji.
- **Błąd standardowy.** Wyświetla błąd standardowy dla każdego oszacowania współczynnika.
- **Przedział ufności.** Wyświetla przedział ufności dla każdego oszacowania współczynnika. Poziom ufności dla przedziału jest ustawiany w oknie dialogowym Opcje.
- **Test T.** Wyświetla test  $t$  każdego oszacowania współczynnika. Hipoteza zerowa dla każdego testu jest taka, że wartość współczynnika wynosi 0.
- **Kowariancje oszacowań parametrów.** Wyświetla oszacowanie macierzy kowariancji dla współczynników modelu.
- **Korelacje oszacowań parametrów.** Wyświetla oszacowanie macierzy korelacji dla współczynników modelu.
- **Efekt planu.** Iloraz wariancji szacowania do wariancji uzyskanej przy założeniu, że pobrana próba jest zwykłą próbą losową. Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.
- **Pierwiastek z efektu planu.** Jest to miara efektu określenia złożonego projektu, w którym wartości dalsze od 1 wskazują na większe efekty.

**Założenia modelu.** Ta grupa umożliwia tworzenie testu założenia proporcjonalnego hazardu. Test porównuje dopasowany model z alternatywnym modelem, który zawiera predyktory zależne od czasu  $x^*_{TF}$  dla każdego predyktora  $x$ , gdzie  $_{TF}$  jest określoną funkcją czasu.

- **Funkcja czasu.** Określa formularz  $_{TF}$  dla modelu alternatywnego. Dla funkcji **identity**:  $_{TF}=T_.$  Dla funkcji **log**:  $_{TF}= \log(T_)$ . Dla **Kaplan-Meier**:  $_{TF}= 1-S_{KM}(T_)$ , gdzie  $S_{KM}(\cdot)$  jest oszacowaniem Kaplana-Meiera dla funkcji przeżycia. W przypadku **rangi**  $_{TF}$  jest porządkiem rangi  $T_$  wśród obserwowanych czasów zakończenia.
- **Oszacowania parametrów dla modelu alternatywnego.** Wyświetla szacowany, standardowy błąd i przedział ufności dla każdego parametru w modelu alternatywnym.
- **Macierz kowariancji dla modelu alternatywnego.** Wyświetla macierz oszacowanych kowariancji między parametrami w modelu alternatywnym.

**Podstawowe funkcje przeżycia i skumulowane funkcje hazardu.** Wyświetla podstawową funkcję przeżycia oraz podstawowe skumulowane hazardy wraz z ich standardowymi błędami.

*Uwaga:* Jeśli predyktory zależne od czasu zdefiniowane na karcie Predyktory zostaną uwzględnione w modelu, ta opcja nie będzie dostępna.

## Wykresy

Karta Wykresy umożliwia wyświetlanie wykresów funkcji zagrożenia, funkcji przeżycia, logarytmiczno-minus-logarytmicznego funkcji przeżycia i jeden minus funkcji przeżycia. Można również wybrać wykres przedziałów ufności wzdłuż określonych funkcji. Poziom ufności jest ustawiany na karcie Opcje.

**Wzorce predyktorów.** Wzorec wartości predyktora, który ma być używany dla żądanych wykresów, oraz eksportowany plik przeżycia można określić na karcie Eksportuj. Należy zauważyć, że opcje te nie są dostępne, jeśli predyktory zależne od czasu zdefiniowane na karcie Predyktory są uwzględniane w modelu.

- **Wykreślanie czynników** Domyślnie każdy czynnik jest oceniany na najwyższym poziomie. W razie potrzeby można wprowadzić lub wybrać inny poziom. Można też wybrać wyświetlenie na wykresie osobnych linii dla każdego poziomu pojedynczego czynnika, zaznaczając pole wyboru dla tego czynnika.
- **Wykreślanie współzmiennych.** Każda współzmienna jest oceniana przy użyciu jej średniej. W razie potrzeby można wprowadzić lub wybrać inną zmienną.

## Testy hipotezy

**Statystyki testu.** Ta grupa umożliwia wybranie typu statystyki używanej do testowania hipotezy. Można wybrać  $F$ , skorygowane  $F$ , chi-kwadrat i skorygowany chi-kwadrat.

**Stopnie swobody w losowaniu.** Ta grupa zapewnia kontrolę nad liczbą stopni swobody projektu losowania, używanych przy obliczaniu wartości  $p$  dla wszystkich statystyk testujących. Wartość bazująca na projekcie losowania to różnica między liczbą jednostek podstawowego losowania a liczbą warstw na pierwszym etapie losowania. Zamiast niej można podać własną liczbę stopni swobody (dodatnią liczbę całkowitą).

**Korekta dla porównań wielokrotnych.** Podczas testowania hipotez z użyciem kontrastów wielokrotnych całościowy poziom istotności można skorygować na podstawie poziomów istotności dla uwzględnionych kontrastów. W tej grupie możliwy jest wybór metody korekty.

- **Najmniejsza istotna różnica.** Ta metoda nie kontroluje ogólnego prawdopodobieństwa odrzucenia hipotez, które stwierdzają, że niektóre kontrasty liniowe różnią się od wartości hipotezy zerowej.
- *Sekwencyjna Sidaka.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sekwencyjna Bonferroniego.* Jest to sekwencyjnie zstępująca odrzucająca procedura Bonferroniego, która jest mniej konserwatywna w zakresie odrzucania indywidualnych hipotez, ale zachowuje identyczny całościowy poziom istotności.
- *Sidak.* Metoda ta zapewnia węższe granice niż test Bonferroniego.
- *Bonferroni.* Metoda ta dostosowuje obserwowany poziom istotności do faktu przeprowadzania wielu porównań.

## Zapisz

**Zapisz zmienne.** Ta grupa pozwala na zapisywanie zmiennych związanych z modelem w aktywnym zbiorze danych w celu dalszego wykorzystania w diagnostyce i do raportowania wyników. Należy pamiętać, że żadna z tych opcji nie jest dostępna, gdy w modelu są uwzględniane predyktory zależne od czasu.

- **Funkcja przeżycia.** Zapisuje prawdopodobieństwo przeżycia (wartość funkcji przeżycia) przy obserwowanych wartościach czasu i predyktora dla każdej obserwacji.
- **Dolna granica przedziału ufności dla funkcji przeżycia.** Zapisuje dolną granicę przedziału ufności dla funkcji przeżycia w obserwowanym czasie i wartościach predyktorów dla każdej obserwacji.
- **Górna granica przedziału ufności dla funkcji przeżycia.** Zapisuje górną granicę przedziału ufności dla funkcji przeżycia w obserwowanym czasie i wartościach predyktorów dla każdej obserwacji.
- **Funkcja skumulowanego hazardu.** Zapisuje skumulowany hazard lub  $-\ln(\text{przeżycie})$  w czasie obserwacji i wartości predyktorów dla każdej obserwacji.
- **Dolna granica przedziału ufności dla funkcji skumulowanego hazardu.** Zapisuje dolną granicę przedziału ufności dla funkcji hazardu w obserwowanym czasie i wartościach predyktorów dla każdej obserwacji.
- **Górna granica przedziału ufności dla funkcji skumulowanego hazardu.** Zapisuje górną granicę przedziału ufności dla funkcji hazardu w obserwowanym czasie i wartościach predyktorów dla każdej obserwacji.
- **Przewidywana wartość predyktora liniowego.** Zapisuje liniową kombinację wartości odniesienia skorygowanych predyktorów pomnożonych przez współczynniki regresji. Predyktor liniowy jest stosunkiem funkcji hazardu do ryzyka bazowego. W przypadku modelu hazardu proporcjonalnego wartość ta jest stała w czasie.
- **Reszta Schoenfelda.** Dla każdej nieobciętej próby i każdego nienadmiarowego parametru w modelu reszta Schoenfelda jest różnicą między obserwowaną wartością predyktora powiązanego z parametrem modelu a oczekiwaną wartością predyktora dla próby w zestawie ryzyka w czasie obserwowanego zdarzenia. Reszty Schoenfelda mogą być używane do oceny założeń dotyczących proporcjonalnego



hazardu; na przykład dla predyktora  $x$  wykresy reszty Schoenfelda dla predyktora zależnego od czasu  $x \cdot \ln(T_0)$  względem czasu powinny pokazywać poziomą linię na poziomie 0, jeśli hazardy są proporcjonalne. Osobna zmienna jest zapisywana dla każdego nienadmiarowego parametru w modelu. Reszty Schoenfelda są obliczane tylko dla przypadków nieobciążonych.

- **Reszta martyngału.** Dla każdego przypadku reszta martyngału jest różnicą między obserwowanym obciążeniem (0, jeśli jest obciążenie, oraz 1 jeśli nie jest), a oczekiwaniem zdarzenia w czasie obserwacji.
- **Reszta odchyień.** Reszt odchyień to reszty martyngału „skorygowane”, aby pojawiać się bardziej symetrycznie wokół wartości 0. Wykresy reszt odchyień względem predyktorów nie powinny ujawniać żadnych wzorców.
- **Reszta Coxa-Snella.** W każdym przypadku reszta Coxa-Snella jest oczekiwaniem zdarzenia w czasie obserwacji lub obserwowanym obciążeniem pomniejszonym o pozostałość martyngału.
- **Reszta wyniku.** Dla każdej obserwacji i każdego nienadmiarowego parametru w modelu, reszta wyniku jest wkładem próby do pierwszej pochodnej pseudowiarygodności. Osobna zmienna jest zapisywana dla każdego nienadmiarowego parametru w modelu.
- **Reszta DFBeta.** Dla każdej próby i każdego nienadmiarowego parametru w modelu reszta DFBeta przybliża zmianę wartości estymacji parametru, gdy próba jest usuwana z modelu. Próby z relatywnie dużymi resztami DFBeta mogą wywierać nadmierny wpływ na analizę. Osobna zmienna jest zapisywana dla każdego nienadmiarowego parametru w modelu.
- **Reszty zagregowane.** Jeśli wiele prób reprezentuje jeden obiekt, zagregowana reszta dla obiektu jest po prostu sumą odpowiednich reszt prób dla wszystkich prób należących do tego samego obiektu. W przypadku reszty Schoenfelda zagregowana wersja jest taka sama, jak wersja niezagregowana, ponieważ reszta Schoenfelda jest definiowana tylko dla prób nieobciążonych. Reszty te są dostępne tylko wtedy, gdy identyfikator obiektu jest określony na karcie Czas i zdarzenie.

**Nazwy zapisywanych zmiennych.** Automatyczne generowanie nazw zapewnia, że praca zostanie zachowana w całości. Nazwy użytkownika umożliwiają likwidację/zastąpienie wyników poprzednich uruchomień bez wcześniejszego usuwania zapisanych zmiennych w Edytorze danych.

## Eksport

**Eksportuj model jako dane IBM SPSS Statistics.** Zapisuje zbiór danych w formacie IBM SPSS Statistics zawierającym korelację parametru lub macierz kowariancji z oszacowaniami parametrów, błędami standardowymi, wartościami istotności oraz stopniami swobody. Kolejność zmiennych w pliku macierzowym jest następująca.

- **rowtype\_.** przyjmuje wartości (i etykiety wartości), COV (kowariancje), CORR (korelacje), EST (oszacowania parametrów), SE (błędy standardowe), SIG (poziomy istotności) i DF (stopnie swobody projektowania próbkowania). Dla każdego parametru modelu istnieje osobna obserwacja z typem wiersza COV (lub CORR), a także osobna obserwacja dla każdego z pozostałych typów wierszy.
- **varname\_.** Przyjmuje wartości P1, P2, ... odpowiadające uporządkowanej liście wszystkich parametrów modelu dla typów wierszy COV lub CORR, z etykietami wartości odpowiadającymi łańcuchom parametrów przedstawionym w tabeli oszacowania parametrów. Te komórki są puste w przypadku wierszy innych typów.
- **P1, P2, ...** Te zmienne odpowiadają uporządkowanej liście wszystkich parametrów modelu z wartościami etykiet odpowiadającymi ciągom znaków parametrów, które są pokazane w tabeli oszacowań parametrów, a ponadto przyjmują wartości zgodnie z typem wiersza. W odniesieniu do parametrów nadmiarowych wszystkie kowariancje są ustawione na zero, korelacje są ustawione na systemowy brak danych; wszystkie oszacowania parametrów są ustawione na zero, a wszystkie błędy standardowe, poziomy istotności i stopnie swobody reszt są ustawione na systemowy brak danych.

*Uwaga:* Tego pliku nie można użyć bezpośrednio do dalszych analiz w innych procedurach, które odczytują plik macierzowy, chyba że te procedury akceptują wszystkie typy wierszy eksportowane w tym przypadku.

**Eksportuj funkcję przeżycia jako dane IBM SPSS Statistics.** Zapisuje zbiór danych w formacie IBM SPSS Statistics zawierającym funkcję przeżycia; standardowy błąd funkcji przeżycia; górne i dolne granice przedziału ufności funkcji przeżycia; a także funkcja skumulowanego hazardu dla każdego niepowodzenia

lub czasu zdarzenia, oceniana dla linii bazowej i dla wzorców predyktorów określonych na karcie Wykres. Kolejność zmiennych w pliku macierzowym jest następująca.

- **Zmienna warstwy podstawowej.** Oddzielne tabele przeżycia są tworzone dla każdej wartości zmiennej warstwy.
- **Zmienna czasu przeżycia.** Czas zdarzenia; dla każdego unikalnego czasu zdarzenia tworzona jest osobna obserwacja.
- **Sur\_0, LCL\_Sur\_0, UCL\_Sur\_0.** Bazowa funkcja przeżycia oraz górna i dolna granica przedziału ufności.
- **Sur\_R, LCL\_Sur\_R, UCL\_Sur\_R.** Funkcja przeżycia wartościowana we wzorcu „odniesienia” (patrz tabela wartości wzorcowych w danych wyjściowych) oraz górna i dolna granica przedziału ufności.
- **Sur\_#.#, LCL\_Sur\_#.#, UCL\_Sur\_#.#, ...** Funkcja przeżycia wartościowana w każdym z wzorców predyktorów określonych na karcie Wykresy oraz górne i dolne granice ich przedziałów ufności. Aby dopasować wzorce z liczbą #.#, należy się zapoznać z tabelą wartości wzorca w danych wyjściowych.
- **Haz\_0, LCL\_Haz\_0, UCL\_Haz\_0.** Bazowa skumulowana funkcja hazardu oraz górna i dolna granica przedziału ufności.
- **Haz\_R, LCL\_Haz\_R, UCL\_Haz\_R.** Skumulowana funkcja hazardu wartościowana we wzorcu „odniesienia” (patrz tabela wartości wzorcowych w danych wyjściowych) oraz górna i dolna granica przedziału ufności.
- **Haz\_#.#, LCL\_Haz\_#.#, UCL\_Haz\_#.#, ...** Skumulowana funkcja hazardu wartościowana w każdym z wzorców predyktorów określonych na karcie Wykresy oraz górne i dolne granice ich przedziałów ufności. Aby dopasować wzorce z liczbą #.#, należy się zapoznać z tabelą wartości wzorca w danych wyjściowych.

**Eksportuj model jako XML.** Zapisuje wszystkie informacje potrzebne do przewidywania funkcji przeżycia, w tym oszacowania parametrów i bazowej funkcji przeżycia w formacie XML (PMML). Możesz użyć tego pliku modelu do stosowania informacji o modelu do innych plików danych w celach statystycznych. .

## Wykresy

**Oszacowanie.** Te elementy sterujące określają kryteria szacowania współczynników regresji.

- **Maksymalna liczba powtórzeń.** Maksymalna liczba iteracji, jakie wykona algorytm. Podaj nieujemną liczbę całkowitą.
- **Maksimum kroków połowienia.** Przy każdej iteracji rozmiar kroku jest zmniejszany o 0,5, aż do momentu wzrostu logarytmu wiarygodności lub osiągnięcia maksymalnej wartości połowienia kroków. Podaj dodatnią liczbę całkowitą.
- **Ogranicz iteracje ze względu na zmiany w ocenach parametrów.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana oszacowania parametru jest mniejsza od określonej wartości, która musi być dodatnia.
- **Ogranicz iteracje z powodu zmian w logarytmie wiarygodności.** Po zaznaczeniu tej opcji algorytm zatrzymuje się, gdy bezwzględna lub względna zmiana funkcji logarytmu wiarygodności jest mniejsza od określonej wartości, która musi być dodatnia.
- **Pokaż przebieg iteracji zgodnie z przyrostem kroków.** Wyświetla przebieg iteracji dla oszacowań parametrów i pseudo logarytmu wiarygodności oraz wyświetla ostatnią ewaluację zmiany wartości szacunkowych parametrów i pseudo logarytmu wiarygodności. Tabela przebiegu iteracji wyświetla wszystkie iteracje  $n$ , począwszy od iteracji 0 (oszacowanie początkowe), gdzie  $n$  jest wartością przyrostu. Jeśli tworzone jest żądanie przebiegu iteracji, wówczas ostatnia iteracja jest zawsze wyświetlana, niezależnie od wartości  $n$ .
- **Metoda przerywania powiązań przy szacowaniu parametrów.** Jeśli istnieją powiązane zaobserwowane czasy niepowodzenia, jedna z tych metod jest używana do przerywania powiązań. Metoda Efrona jest bardziej kosztowna obliczeniowo.

**Funkcje przeżycia.** Te elementy sterujące określają kryteria dla obliczeń z udziałem funkcji przeżycia.

- **Metoda szacowania funkcji podstawowego przeżycia.** Metoda **Breslowa** (lub Nelsona-Aalena albo empiryczna) szacuje podstawowy skumulowany hazard za pomocą niemalejącej funkcji krokowej

z krokami w obserwowanych czasach niepowodzenia, a następnie oblicza podstawowe przeżycie za pomocą zależności  $\text{przeżycie} = \exp(-\text{skumulowany hazard})$ . Metoda **Efrona** jest bardziej kosztowna obliczeniowo i redukuje się do metody Breslowa, gdy nie ma żadnych powiązań. Metoda **limitu produktu** szacuje początkowe przeżycie za pomocą nierosnącej funkcji prawostronnie ciągłej; gdy w modelu nie ma predyktorów, metoda ta sprowadza się do szacowania Kaplana-Meiera.

- **Przedziały ufności dla funkcji przeżycia.** Przedział ufności można obliczyć na trzy sposoby: w oryginalnych jednostkach, za pośrednictwem transformacji logarytmicznej lub transformacji logarytm minus logarytm. Tylko transformacja logarytm minus logarytm gwarantuje, że granice przedziału ufności będą znajdować się w zakresie od 0 do 1, ale transformacja logarytmiczna zwykle wydaje się sprawdzać „najlepiej”.

**Brakujące wartości użytkownika** Wszystkie zmienne muszą mieć poprawne wartości dla obserwacji, która ma zostać uwzględniona w analizie. Te elementy sterujące umożliwiają określenie, czy brakujące wartości użytkownika są traktowane jako poprawne dla modeli jakościowych (w tym czynników, zdarzeń, warstw i zmiennych podpopulacji), a także zmiennych projektu próbkowania.

**Przedział ufności (%).** Jest to poziom przedziału ufności używany dla oszacowań współczynników, oszacowań współczynników wykładniczych, oszacowań funkcji przeżycia i oszacowań funkcji skumulowanego hazardu. Podaj wartość większą lub równą 0 i mniejszą niż 100.

## Dodatkowe właściwości komendy CSCOXREG

Język składni komend umożliwia również:

- Wykonanie testów hipotez użytkownika (przy użyciu opcji komendy CUSTOM i komendy /PRINT LMATRIX).
- Określenie tolerancji (przy użyciu komendy /CRITERIA SINGULAR).
- Uzyskanie dostępu do tabeli ogólnych funkcji estymowalnych (przy użyciu komendy /PRINT GEF).
- Uzyskanie dostępu do wielu wzorców predyktorów (przy użyciu wielu opcji komendy PATTERN).
- Uzyskanie informacji o maksymalnej liczbie zapisanych zmiennych po podaniu trzonu nazwy (przy użyciu opcji komendy SAVE). W oknie dialogowym CSCOXREG domyślnie uwzględnianych jest 25 zmiennych.

Pełne informacje na temat składni znajdują się w podręczniku *Command Syntax Reference*.



## Uwagi

---

Niniejsza publikacja została przygotowana z myślą o produktach i usługach oferowanych w Stanach Zjednoczonych. IBM może udostępniać ten materiał w innych językach. Jednakże w celu uzyskania dostępu do takiego materiału istnieje konieczność posiadania egzemplarza produktu w takim języku.

Produktów, usług lub opcji opisywanych w tym dokumencie IBM nie musi oferować we wszystkich krajach. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela IBM. Odwołanie do produktu, programu lub usługi IBM nie oznacza, że można użyć wyłącznie tego produktu, programu lub usługi IBM. Zamiast nich można zastosować ich odpowiednik funkcjonalny pod warunkiem że nie narusza to praw własności intelektualnej IBM. Jednakże cała odpowiedzialność za ocenę przydatności i sprawdzenie działania produktu, programu lub usługi pochodzących od producenta innego niż IBM spoczywa na użytkowniku.

IBM może posiadać patenty lub złożone wnioski patentowe na produkty, o których mowa w niniejszej publikacji. Przedstawienie tej publikacji nie daje żadnych uprawnień licencyjnych do tychże patentów. Pisemne zapytania w sprawie licencji można przesyłać na adres:

*IBM Director of Licensing*

*IBM Corporation*

*North Castle Drive, MD-NC119  
Armonk, NY 10504-1785 U.S.A.*

Zapytania dotyczące zestawów znaków dwubajtowych (DBCS) należy kierować do lokalnych działów własności intelektualnej IBM (IBM Intellectual Property Department) lub wysłać je na piśmie na adres:

*Intellectual Property Licensing*

*Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokio 103-8510, Japonia*

INTERNATIONAL BUSINESS MACHINES CORPORATION DOSTARCZA TĘ PUBLIKACJĘ W STANIE, W JAKIM SIĘ ZNAJDUJE ("AS IS") BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (W TYM TAKŻE RĘKOJMI), WYRAŻNYCH LUB DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU ORAZ GWARANCJI, ŻE PUBLIKACJA NIE NARUSZA PRAW STRON TRZECICH. Ustawodawstwa niektórych krajów nie dopuszczają zastrzeżeń dotyczących gwarancji wyraźnych lub domniemanych w odniesieniu do pewnych transakcji; w takiej sytuacji powyższe zdanie nie ma zastosowania.

Informacje zawarte w tej publikacji mogą zawierać nieścisłości techniczne lub błędy drukarskie. Informacje te są okresowo aktualizowane, a zmiany te zostaną uwzględnione w kolejnych wydaniach tej publikacji. IBM zastrzega sobie prawo do wprowadzania ulepszeń i/lub zmian w produktach i/lub programach opisanych w tej publikacji w dowolnym czasie, bez wcześniejszego powiadomienia.

Wszelkie wzmianki w tej publikacji na temat stron internetowych firm innych niż IBM zostały wprowadzone wyłącznie dla wygody użytkowników i w żadnym razie nie stanowią zachęty do ich odwiedzania. Materiały dostępne na tych stronach nie są częścią materiałów opracowanych dla tego produktu IBM, a użytkownik korzysta z nich na własną odpowiedzialność.

IBM ma prawo do używania i rozpowszechniania informacji przystanych przez użytkownika w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich autora.

Licencjodawcy tego programu, którzy chcieliby uzyskać informacje na temat programu w celu: (i) umożliwienia wymiany informacji między niezależnie utworzonymi programami i innymi programami

(łącznie z opisywanym) oraz (ii) wykorzystywania wymienianych informacji, powinni skontaktować się z:

*IBM Director of Licensing*

*IBM Corporation*

*North Castle Drive, MD-NC119  
Armonk, NY 10504-1785 U.S.A.*

Informacje takie mogą być udostępnione, o ile spełnione zostaną odpowiednie warunki, w tym, w niektórych przypadkach, zostanie uiszczona stosowna opłata.

Licencjonowany program opisany w niniejszej publikacji oraz wszystkie inne licencjonowane materiały dostępne dla tego programu są dostarczane przez IBM na warunkach określonych w Umowie IBM z Klientem, Międzynarodowej Umowie Licencyjnej IBM na Program lub w innych podobnych umowach zawartych między IBM i użytkownikami.

Dane dotyczące wydajności i cytowane przykłady zostały przedstawione jedynie w celu zobrazowania sytuacji. Faktyczne wyniki dotyczące wydajności mogą się różnić w zależności do konkretnych warunków konfiguracyjnych i operacyjnych.

Informacje dotyczące produktów innych podmiotów niż IBM zostały uzyskane od dostawców tych produktów, z ich publicznych ogłoszeń lub innych dostępnych publicznie źródeł. IBM nie testował tych produktów i nie może potwierdzić dokładności pomiarów wydajności, kompatybilności ani żadnych innych danych związanych z produktami firm innych niż IBM. Pytania dotyczące możliwości produktów firm innych niż IBM należy kierować do dostawców tych produktów.

Wszelkie stwierdzenia dotyczące przyszłych kierunków rozwoju i zamierzeń IBM mogą zostać zmienione lub wycofane bez powiadomienia.

Publikacja ta zawiera przykładowe dane i raporty używane w codziennych operacjach działalności gospodarczej. W celu kompleksowego zilustrowania tej działalności podane przykłady zawierają nazwy osób, firm i ich produktów. Wszystkie te nazwy/nazwiska są fikcyjne i jakiegokolwiek podobieństwo do istniejących nazw/nazwisk jest całkowicie przypadkowe.

#### LICENCJA W ZAKRESIE PRAW AUTORSKICH:

Niniejsza publikacja zawiera przykładowe aplikacje w kodzie źródłowym ilustrujące techniki programowania w różnych systemach operacyjnych. Użytkownik może kopiować, modyfikować i rozpowszechniać te programy przykładowe w dowolnej formie bez uiszczania opłat na rzecz IBM, w celu rozbudowy, użytkowania, handlowego lub w celu rozpowszechniania aplikacji zgodnych z aplikacyjnym interfejsem programowym dla tego systemu operacyjnego, dla którego napisane były programy przykładowe. Programy przykładowe nie zostały gruntownie przetestowane. IBM nie może zatem gwarantować ani sugerować niezawodności, użyteczności i funkcjonalności tych programów. Programy przykładowe są dostarczane w stanie, w jakim się znajdują ("AS IS"), bez jakichkolwiek gwarancji (rękojmię również wyłącza się). IBM nie ponosi odpowiedzialności za jakiegokolwiek szkody wynikające z używania programów przykładowych.

Każda kopia programu przykładowego lub jakiegokolwiek jego fragment, jak też jakiegokolwiek prace pochodne muszą zawierać następujące uwagi dotyczące praw autorskich:

© Copyright IBM Corp. 2021. Fragmenty tego kodu pochodzą z przykładowych programów produktu IBM Corp. Programy przykładowe.

© Copyright IBM Corp. 1989-2021. Wszelkie prawa zastrzeżone.

## Znaki towarowe

IBM, logo IBM i ibm.com są znakami towarowymi lub zastrzeżonymi znakami towarowymi International Business Machines Corp., zarejestrowanymi w wielu systemach prawnych na całym świecie. Pozostałe nazwy produktów i usług mogą być znakami towarowymi IBM lub innych przedsiębiorstw. Aktualna lista znaków towarowych IBM dostępna jest w serwisie WWW, w sekcji "Copyright and trademark

information" (Informacje o prawach autorskich i znakach towarowych), pod adresem [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Adobe, logo Adobe, PostScript oraz logo PostScript są znakami towarowymi lub zastrzeżonymi znakami towarowymi Adobe Systems Incorporated w Stanach Zjednoczonych i/lub w innych krajach.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium i Pentium są znakami towarowymi lub zastrzeżonymi znakami towarowymi Intel Corporation lub przedsiębiorstw podporządkowanych w Stanach Zjednoczonych i w innych krajach.

Linux jest zastrzeżonym znakiem towarowym Linusa Torvaldsa w Stanach Zjednoczonych i/lub w innych krajach.

Microsoft, Windows, Windows NT oraz logo Windows są znakami towarowymi Microsoft Corporation w Stanach Zjednoczonych i/lub w innych krajach.

UNIX jest zastrzeżonym znakiem towarowym Open Group w Stanach Zjednoczonych i w innych krajach.

Java oraz wszystkie znaki towarowe i logo dotyczące Java są znakami towarowymi firmy i jej firm zależnych.





# Indeks

## B

- błąd standardowy
  - w częstościach dla prób złożonych [12](#)
  - w ilorazach dla prób złożonych [17](#)
  - w ogólnym modelu liniowym dla prób złożonych [19](#)
  - w regresji logistycznej dla prób złożonych [23](#)
  - w regresji porządkowej prób złożonych [28](#)
  - w statystykach opisowych dla prób złożonych [14](#)
  - w tabelach krzyżowych dla prób złożonych [15](#)
- Bonferroniego
  - w próbach złożonych [20](#), [24](#), [29](#)
  - w przypadku regresji Coxa dla prób złożonych [36](#)
- braki danych
  - w regresji logistycznej dla prób złożonych [25](#)
- Brakujące dane
  - w statystykach ilorazowych dla prób złożonych [17](#)
- brakujące wartości
  - w ogólnym modelu liniowym dla prób złożonych [21](#)
  - w próbach złożonych [12](#), [16](#)
  - w regresji porządkowej prób złożonych [30](#)
  - w statystykach opisowych dla prób złożonych [14](#)

## C

- chi-kwadrat
  - w próbach złożonych [20](#), [24](#), [29](#)
  - w przypadku regresji Coxa dla prób złożonych [36](#)
- Częstości dla prób złożonych statystyki [12](#)

## D

- dobór próby
  - projekt złożony [2](#)

## E

- efekt planu
  - w częstościach dla prób złożonych [12](#)
  - w ilorazach dla prób złożonych [17](#)
  - w ogólnym modelu liniowym dla prób złożonych [19](#)
  - w przypadku regresji Coxa dla prób złożonych [34](#)
  - w regresji logistycznej dla prób złożonych [23](#)
  - w regresji porządkowej prób złożonych [28](#)
  - w statystykach opisowych dla prób złożonych [14](#)
  - w tabelach krzyżowych dla prób złożonych [15](#)

## I

- ilorazy szans
  - w regresji logistycznej dla prób złożonych [24](#)
  - w regresji porządkowej prób złożonych [29](#)
  - w tabelach krzyżowych dla prób złożonych [15](#)
- informacja o planie losowania
  - w przypadku regresji Coxa dla prób złożonych [34](#)

## iteracje

- w regresji logistycznej dla prób złożonych [25](#)
- w regresji porządkowej prób złożonych [30](#)

## K

- kategoria odniesienia
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
  - w regresji logistycznej dla prób złożonych [22](#)
- klastry
  - w kreatorze losowania prób [3](#)
- kontrasty
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- kontrasty Helmerta
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- kontrasty odchylenia
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- kontrasty powtarzane
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- kontrasty różnicy
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- kontrasty wielomianowe
  - w ogólnym modelu liniowym dla prób złożonych [20](#)
- korelacje oszacowań parametrów
  - w ogólnym modelu liniowym dla prób złożonych [19](#)
  - w regresji logistycznej dla prób złożonych [23](#)
  - w regresji porządkowej prób złożonych [28](#)
- kowariancje oszacowań parametrów
  - w ogólnym modelu liniowym dla prób złożonych [19](#)
  - w regresji logistycznej dla prób złożonych [23](#)
  - w regresji porządkowej prób złożonych [28](#)
- kroki połowienia
  - w regresji logistycznej dla prób złożonych [25](#)
  - w regresji porządkowej prób złożonych [30](#)

## L

- losowanie proste
  - w kreatorze losowania [4](#)
- losowanie prób
  - w kreatorze losowania [4](#)
- losowanie prób PPS
  - w kreatorze losowania [4](#)

## M

- metoda losowania
  - w kreatorze losowania [4](#)
- metoda losowania prób Brewera
  - w kreatorze losowania [4](#)
- Metoda losowania prób Murthy'ego
  - w kreatorze losowania [4](#)
- Metoda losowania Sampforda
  - w kreatorze losowania [4](#)
- metoda Newtona-Raphsona
  - w regresji porządkowej prób złożonych [30](#)

Metoda szacowania Breslowa  
w przypadku regresji Coxa dla prób złożonych [38](#)  
Metoda szacowania Efrona  
w przypadku regresji Coxa dla prób złożonych [38](#)  
miara wielkości  
w kreatorze losowania [4](#)

## N

najmniejsza istotna różnica  
w próbach złożonych [20](#), [24](#), [29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)  
nieważona liczebność  
w częstościach dla prób złożonych [12](#)  
w ilorazach dla prób złożonych [17](#)  
w statystykach opisowych dla prób złożonych [14](#)  
w tabelach krzyżowych dla prób złożonych [15](#)

## O

Ocena Fishera  
w regresji porządkowej prób złożonych [30](#)  
Ogólny model liniowy dla prób złożonych  
dodatkowe właściwości komendy [21](#)  
model [19](#)  
opcje [21](#)  
oszacowane średnie [20](#)  
statystyki [19](#)  
zapisywanie zmiennych [21](#)  
oszacowania parametrów  
w ogólnym modelu liniowym dla prób złożonych [19](#)  
w przypadku regresji Coxa dla prób złożonych [34](#)  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)

## P

pierwiastek z efektu planu  
w częstościach dla prób złożonych [12](#)  
w ilorazach dla prób złożonych [17](#)  
w ogólnym modelu liniowym dla prób złożonych [19](#)  
w przypadku regresji Coxa dla prób złożonych [34](#)  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)  
w statystykach opisowych dla prób złożonych [14](#)  
w tabelach krzyżowych dla prób złożonych [15](#)  
plan analizy [8](#)  
plan próbkowania [2](#)  
plik planu [2](#)  
podpopulacja  
w przypadku regresji Coxa dla prób złożonych [34](#)  
podział na warstwy  
w kreatorze losowania prób [3](#)  
w kreatorze przygotowania analizy [8](#)  
poprawka Sidaka  
w próbach złożonych [20](#), [24](#), [29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)  
poziom ufności  
w regresji logistycznej dla prób złożonych [25](#)  
w regresji porządkowej prób złożonych [30](#)  
prawdopodobieństwa odpowiedzi  
w regresji porządkowej prób złożonych [27](#)  
prawdopodobieństwa włączenia

prawdopodobieństwa włączenia (*kontynuacja*)  
w kreatorze losowania prób [5](#)  
predyktor zależny od czasu  
w przypadku regresji Coxa dla prób złożonych [33](#)  
procent w kolumnie  
w tabelach krzyżowych dla prób złożonych [15](#)  
procent w wierszu  
w tabelach krzyżowych dla prób złożonych [15](#)  
procenty w tabeli  
w częstościach dla prób złożonych [12](#)  
w tabelach krzyżowych dla prób złożonych [15](#)  
proporcja próby  
w kreatorze losowania prób [5](#)  
proste kontrasty  
w ogólnym modelu liniowym dla prób złożonych [20](#)  
próbkowanie złożone  
plan analizy [8](#)  
plan próbkowania [2](#)  
Próby złożone  
brakujące wartości [12](#), [16](#)  
opcje [13](#), [14](#), [16](#), [18](#)  
testy hipotez [20](#), [24](#), [29](#)  
Próby złożone - Statystyki opisowe  
brakujące wartości [14](#)  
statystyki [14](#)  
przebieg iteracji  
w regresji logistycznej dla prób złożonych [25](#)  
w regresji porządkowej prób złożonych [30](#)  
przedziały ufności  
w częstościach dla prób złożonych [12](#)  
w ilorazach dla prób złożonych [17](#)  
w ogólnym modelu liniowym dla prób złożonych [19](#), [21](#)  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)  
w statystykach opisowych dla prób złożonych [14](#)  
w tabelach krzyżowych dla prób złożonych [15](#)  
przewidywane kategorie  
w regresji logistycznej dla prób złożonych [25](#)  
w regresji porządkowej prób złożonych [29](#)  
przewidywane prawdopodobieństwo  
w regresji logistycznej dla prób złożonych [25](#)  
w regresji porządkowej prób złożonych [29](#)

## R

Regresja Coxa dla prób złożonych  
definiowanie zdarzenia [33](#)  
eksport modelu [37](#)  
model [34](#)  
opcje [38](#)  
podgrupy [34](#)  
predyktor zależny od czasu [33](#)  
predyktory [33](#)  
statystyki [34](#)  
testy hipotez [36](#)  
wykresy [35](#)  
zapisywanie zmiennych [36](#)  
regresja Coxa prób złożonych  
Analiza Kaplana-Meiera [31](#)  
zmiennie daty i czasu [31](#)  
Regresja logistyczna dla prób złożonych  
dodatkowe właściwości komendy [26](#)  
ilorazy szans [24](#)  
kategoria odniesienia [22](#)

Regresja logistyczna dla prób złożonych (*kontynuacja*)  
model [23](#)  
opcje [25](#)  
statystyki [23](#)  
zapisywanie zmiennych [25](#)

Regresja porządkowa prób złożonych  
ilorazy szans [29](#)  
model [27](#)  
opcje [30](#)  
prawdopodobieństwa odpowiedzi [27](#)  
statystyki [28](#)  
zapisywanie zmiennych [29](#)

reszty  
w ogólnym modelu liniowym dla prób złożonych [21](#)  
w tabelach krzyżowych dla prób złożonych [15](#)

reszty Coxa-Snella  
w przypadku regresji Coxa dla prób złożonych [36](#)

reszty częściowe Schoenfelda  
w przypadku regresji Coxa dla prób złożonych [36](#)

reszty martynałów  
w przypadku regresji Coxa dla prób złożonych [36](#)

reszty odchylenia  
w przypadku regresji Coxa dla prób złożonych [36](#)

reszty skorygowane  
w tabelach krzyżowych dla prób złożonych [15](#)

reszty wyniku  
w przypadku regresji Coxa dla prób złożonych [36](#)

reszty zagregowane  
w przypadku regresji Coxa dla prób złożonych [36](#)

różnica ryzyka  
w tabelach krzyżowych dla prób złożonych [15](#)

ryzyko względne  
w tabelach krzyżowych dla prób złożonych [15](#)

## S

sekwencyjna poprawka Bonferroniego  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

sekwencyjna poprawka Sidaka  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

sekwencyjne losowanie  
w kreatorze losowania [4](#)

separacja  
w regresji logistycznej dla prób złożonych [25](#)  
w regresji porządkowej prób złożonych [30](#)

skorygowana statystyka F  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

skorygowany Chi-kwadrat  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

skumulowane prawdopodobieństwa  
w regresji porządkowej prób złożonych [29](#)

skupienia  
w kreatorze przygotowania analizy [8](#)

statystyka R2  
w ogólnym modelu liniowym dla prób złożonych [19](#)

statystyki F  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

Statystyki ilorazowe dla prób złożonych  
Brakujące dane [17](#)

Statystyki ilorazowe dla prób złożonych (*kontynuacja*)  
statystyki [17](#)

statystyki pseudo R kwadrat  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)

stopnie swobody  
w próbach złożonych [20, 24, 29](#)  
w przypadku regresji Coxa dla prób złożonych [36](#)

suma  
w statystykach opisowych dla prób złożonych [14](#)

szacowane średnie brzegowe  
w ogólnym modelu liniowym dla prób złożonych [20](#)

szacowanie losowania  
w kreatorze przygotowania analizy [9](#)

## Ś

średnia  
w statystykach opisowych dla prób złożonych [14](#)

## T

tabele klasyfikacji  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)

Tabele krzyżowe dla prób złożonych  
statystyki [15](#)

Test linii równoległych  
w regresji porządkowej prób złożonych [28](#)

test proporcjonalnych hazardów  
w przypadku regresji Coxa dla prób złożonych [34](#)

test t  
w ogólnym modelu liniowym dla prób złożonych [19](#)  
w regresji logistycznej dla prób złożonych [23](#)  
w regresji porządkowej prób złożonych [28](#)

## W

wagi próby  
w kreatorze losowania prób [5](#)  
w kreatorze przygotowania analizy [8](#)

warstwa podstawowa  
w przypadku regresji Coxa dla prób złożonych [34](#)

wartości oczekiwane  
w tabelach krzyżowych dla prób złożonych [15](#)

wartości przewidywane  
w ogólnym modelu liniowym dla prób złożonych [21](#)

wartości skumulowane  
w częstościach dla prób złożonych [12](#)

wejściowe wagi losowania  
w kreatorze losowania prób [3](#)

wielkość populacji  
w częstościach dla prób złożonych [12](#)  
w ilorazach dla prób złożonych [17](#)  
w kreatorze losowania prób [5](#)  
w statystykach opisowych dla prób złożonych [14](#)  
w tabelach krzyżowych dla prób złożonych [15](#)

wielkość próby  
w kreatorze losowania prób [5](#)  
w kreatorze próbkowania [4](#)

współczynnik zmienności (COV)  
w częstościach dla prób złożonych [12](#)  
w ilorazach dla prób złożonych [17](#)

współczynnik zmienności (COV) (*kontynuacja*)  
w statystykach opisowych dla prób złożonych 14  
w tabelach krzyżowych dla prób złożonych 15

## Z

zbieżność parametru  
w regresji logistycznej dla prób złożonych 25  
w regresji porządkowej prób złożonych 30

zbieżność wiarygodności  
w regresji logistycznej dla prób złożonych 25  
w regresji porządkowej prób złożonych 30



