**Temat projektu:** Modele nieparametryczne w analizie historii zdarzeń

Autor: Sebastian Boruch

Spis treści

[1. Wprowadzenie 2](#_Toc507522009)

[2. Opis danych 2](#_Toc507522010)

[2.1. Źródło 2](#_Toc507522011)

[2.2. Cechy zbioru 3](#_Toc507522012)

[3. Estymacja modeli nieparametrycznych 6](#_Toc507522013)

[3.1. Metoda tradycyjna 6](#_Toc507522014)

[3.2. Metoda tradycyjna ze stratyfikacją 9](#_Toc507522015)

[3.3. Metoda Kaplana-Meiera 14](#_Toc507522016)

[4. Porównanie modeli i wnioski 16](#_Toc507522017)

[5. Bibliografia 16](#_Toc507522018)

[6. Kody SAS 16](#_Toc507522019)

# 1. Wprowadzenie

Leczenie narkomanów to jeden z kluczowych problemów medycyny w ostatnich dekadach. Rozwój technologii, osamotnienie czy nadmiar obowiązków sprzyja rozwojowi czynników stresogennych, co z kolei prowadzi do szukania substancji uśmierzających dyskomfort psychiczny jak np. narkotyki. Wzrost dochodów per capita i wzrastająca dostępność środków odurzających (np. poprzez zakup w internecie) powoduje, że ryzyko zachorowania rośnie. W poniższym opracowaniu skupię się na czynnikach pozwalających na wyleczenie uzależnionego. W pierwszej części przedstawię dane, na których przeprowadzę analizę- interpretacja i testowanie wszelkich dostępnych testów istotności oraz testu homogeniczności funkcji dożycia w warstwach, opis zbioru danych, analizę zmiennych i definicję modelowanego zjawiska. W drugiej części stworzę tablice przetrwania metodą tradycyjną i Kaplana-Meiera za pomocą procedury LIFETEST systemu SAS.

# 2. Opis danych

## 2.1. Źródło

Australijskie badanie przeprowadzone przez Caplehorna i Bell (1991) z departamentu zdrowia publicznego Uniwersytetu w Sydney porównywało czas przeżycia w dwóch klinikach leczenia metadonem dla osób uzależnionych od heroiny. Czas przeżycia pacjenta określono jako czas w dniach, aż pacjent powrócił do nałogu. Obie kliniki różniły się w zależności od ogólnej polityki leczenia. Celem było zidentyfikowanie czynników, które wpływają na czas przeżycia: klinika, maksymalna dzienna dawka metadonu i (nie)obecność w więzieniu.

Proces leczenia pacjentów metadonem badano w kohorcie 238 osób uzależnionych od heroiny, którzy weszli do programu terapii uzależnień od lutego 1986 do sierpnia 1987 r. Wszyscy pacjenci zostali ocenieni w tym samym instytucie i odesłani do jednej z dwóch placówek leczniczych w celu rozpoczęcia terapii.

Czym jest metadon?

Metadon wykorzystywany jest w leczeniu osób uzależnionych od opiatów, m.in. heroiny i morfiny. Substancja znosi objawy abstynencyjne, ale jednocześnie sama posiada duży potencjał uzależniający, dlatego jej stosowanie powinno odbywać się pod ścisłą kontrolą lekarską. W razie przedawkowania metadonu pacjent naraża się na kolejne uzależnienie, którego skutki mogą być dużo poważniejsze niż w przypadku zażywania heroiny. Dlatego indywidualny dobór dawki jest tu kluczowy dla wyleczenia uzależnionego.

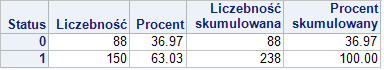
## 2.2. Cechy zbioru

Początki badania sięgają lutego 1986, lecz nie jest podana dokłada data zakończenia zbierania danych. Biorąc pod uwagę fakt, że najdłuższy okres przeżycia sięga 1100 dni, a wyniki badań opublikowane zostały w roku 1991 autor ocenia, że okres zbierania danych do analizy zakończył się w roku 1990. W zbiorze nie występują braki danych.

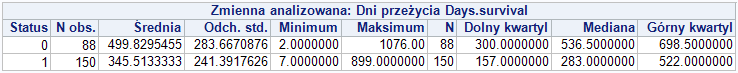
|  |  |  |  |
| --- | --- | --- | --- |
| Nazwa zmiennej | Opis | Typ | Wartości |
| Klinika | Placówka, w której pacjent był leczony | Jakościowa | 1- klinika nr 1  2- klinika nr 2 |
| Status | Przyczyna zakończenia naliczania dni pobytu pacjenta w klinice | Binarna | 0- pacjent opuścił klinikę (cenzura prawostronna)  1- powrót do nałogu |
| Dni przeżycia | Liczba dni pobytu w klinice | Ilościowa ciągła | od 2 do 1076 |
| Więzienie | Czy pacjent przebywał kiedykolwiek w więzieniu? | Binarna | 0- nie  1- tak |
| Dawka | Maksymalna dzienna dawka metadonu (mg/dzień) | ilościowa | 20-110 |

*Tabela 1. charakterystyka zmiennych*

Zdarzeniem (porażką) jest tu powrót do nałogu przez pacjenta (status=1). Cenzorowanie następuje gdy zmienna status=0. Porażka wystąpiła w 63% przypadków:



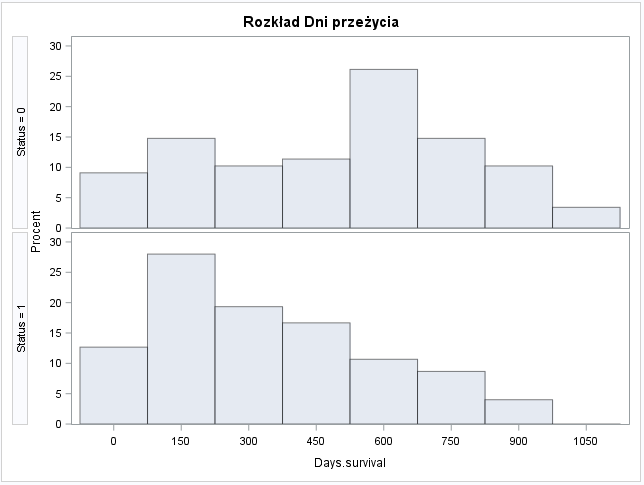
*Tabela 2. Rozkład obserwacji, dla zmiennej status*



*Tabela 3. Miary statystyczne zmiennej czasu z podziałem na porażkę lub jej brak*

Jeżeli zaszło zdarzenie, to przeciętny czas przeżycia pacjenta wynosił 345 dni. Mediana to tylko 283, więc rozkład jest prawostronnie asymetryczny.

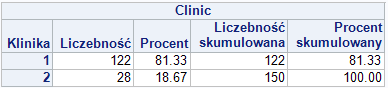
Jeżeli do zdarzenia nie doszło, to średni czas przeżycia pacjenta w klinice wyniósł 500 dni i jest to mniej niż mediana (537) więc rozkład jest lewostronnie asymetryczny.



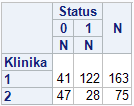
*Wykres 1. Rozkład dni przeżycia na podstawie zmiennej STATUS*

Z wykresu można wnioskować, że osoby które wracają do nałogu, czynią to przeważnie po około pół roku pobytu w klinice. Osoby, które do nałogu nie powracają, muszą przeważnie przebywać w kliniec około półtora roku.

Poza zmienną STATUS występują również 3 inne zmienne objaśniające model. Jedną z nich jest zmienna KLINIKA.

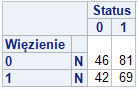


*Tabela 4. Rozkład dla zmiennej KLINIKA, gdzie zaszło zdarzenie*



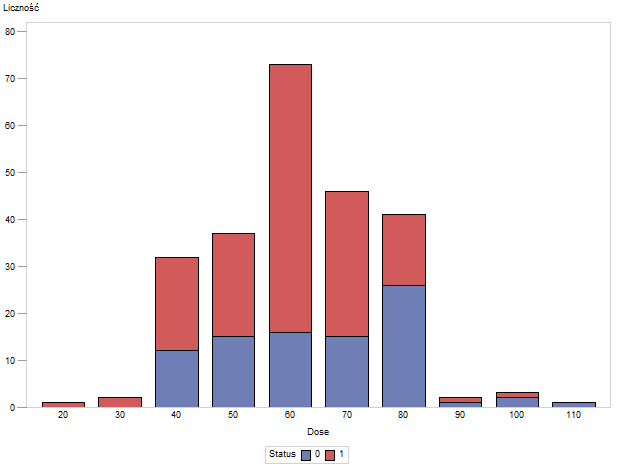
*Tabela 5. Zagregowana tabela liczebności zmiennych KLINIKA i STATUS*

Powyższe tabele sugerują, że zmienna KLINIKA jest istotna statystycznie. Ryzyko zdarzenia znacząco wzrasta przy pobycie pacjenta w klinice 1. Odsetek pacjentów, którzy wrócili do nałogu w klinice 2. jest znacząco niższy.



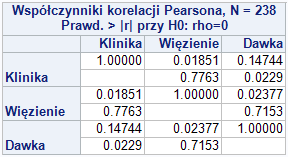
*Tabela 6. Zagregowana tabela liczebności zmiennych WIĘZIENIE i STATUS*

Kolejną zmienną biorącą udział w badaniu była zmienna WIĘZIENIE. Dostarcza ona informacji czy dany pacjent kiedykolwiek przebywał z zakładzie karnym. Po analizie powyższej tabeli można zasugerować wniosek, iż pobyt w więzieniu nie wpływa istotnie na ryzyko powrotu do nałogu. Jeżeli jednak należałoby wyciągnąć jakieś wnioski, to można stwierdzić, że pacjenci którzy przebywali wcześniej w więzieniu, mają mniejszą skłonność do powrotu do nałogu.



*Wykres 2. Rozkład zmiennej DAWKA z uwzględnieniem zmiennej STATUS*

Ostatnią zmienną uwzględnioną w badaniu była zmienna DAWKA, która informowała o przyjmowanej przez pacjenta dawce metadonu. Po analizie powyższego wykresu można stwierdzić, że odsetek porażki przewyższa znacząco odsetek przeżycia do dawki 70 mg/dzień. Powyżej tej dawki przeważa ilość pacjentów, którzy nie powrócili do nałogu. Rozkład zmiennej jest zbliżony do rozkładu normalnego.



*Tabela 7. Tabela korelacji zmiennych*

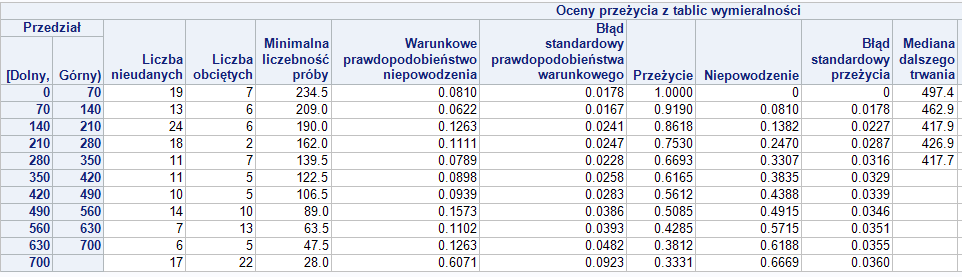
Powyższa tabela korelacji ukazuje, że nie występuje silna korelacja pomiędzy żadną ze zmiennych. Co za tym idzie, nie trzeba usuwać żadnej ze zmiennych ani dodawać nowych. Nie trzeba również dodawać efektów interakcji.

# 3. Estymacja modeli nieparametrycznych

W niniejszym rozdziale przedstawiono estymację modeli nieparametrycznych i konstrukcję tablic trwania życia przy użyciu kolejno metody tradycyjnej z użyciem stratyfikacji oraz metody Kaplana-Meiera. Kroków tych dokonano za pomocą procedury PROC LIFETEST.

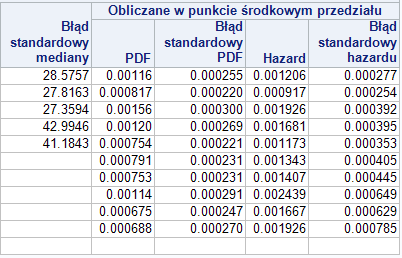
## 3.1. Metoda tradycyjna

Tradycyjne tablice trwania życia są metodą, która dostarcza wyników dla funkcji dożycia, funkcji gęstości i wskaźnika hazardu. Jednym z wymogów tej metody jest uzyskanie takiej struktury danych, aby zmienna czasowa miała przedziały o podobnej długości. Zmienną czasową jest tu zmienna DNI PRZEŻYCIA a zmienną cenzurującą zmienna STATUS. Jeżeli zmienna STATUS przyjmie wartość 0- obserwacja jest obcięta, jeżeli wartość 1- nieudana (zachodzi porażka).



*Tabela 8. Podstawowe wyniki estymacji modelu nieparametrycznego metodą tradycyjną.*

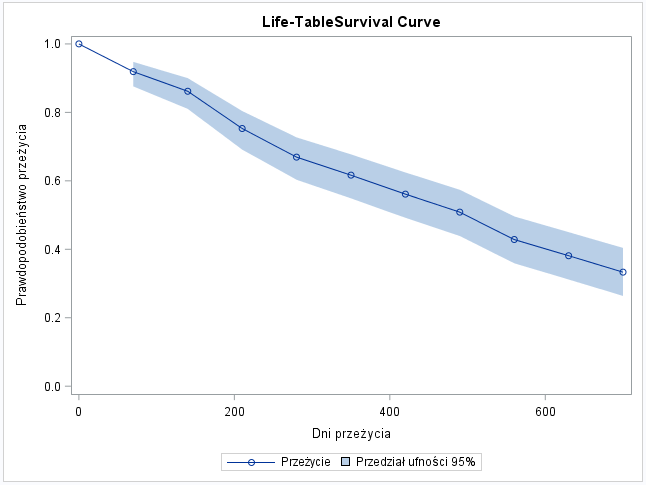
Zmienna czasu została podzielona na 70-dniowe interwały z górną granicą 700 dni. Każdy z przedziałów ma satysfakcjonującą dużą liczebność i wystarczająco mały odsetek danych odciętych bądź nieudanych, co pozwala przeprowadzić analizę.



*Tabela 9. Podstawowe wyniki estymacji obliczane w punkcie środkowym przedziału*

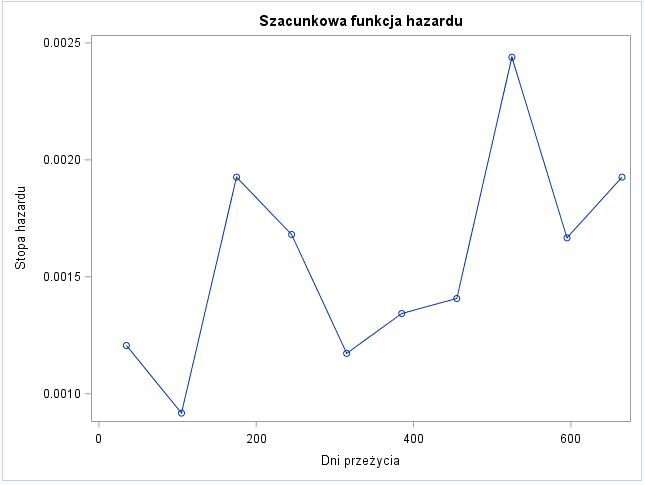
Autor podjął się analizy przedziału, gdzie występuje najwięcej porażek (140-210):

* liczba nieudanych: **24** - porażki, czyli liczba pacjentów, która powróciła do nałogu w tym przedziale czasu
* liczba obciętych: **6** - liczba obserwacji ocenzurowanych- pacjent opuścił klinikę
* minimalna liczebność próby- wielkość efektywna próby w tym przedziale to 190:
  + metoda obliczenia: 209-13- ½\*6- ½\*6 = 190
  + Całkowita liczebność próby (populacja) – suma z kolumny liczba nieudanych do przedziału poprzedzającego włącznie – suma z kolumny liczba obciętych do przedziału poprzedzającego włącznie -1/2\* liczba obciętych dla badanego przedziału.
* Warunkowe prawdopodobieństwo niepowodzenia: **0,1263** – warunkowe prawdopodobieństwo zajścia zdarzenia, a zatem powrotu do nałogu dla badanego przedziału wynosi 12,63%, zatem prawdopodobieństwo dożycia 210. dnia (koniec przedziału) wynosi 100%−12,63%=87,37%
* Przeżycie: **0,8618** prawdopodobieństwo, że osoba nie wróci do nałogu do początku badanego przedziału (prawdopodobieństwo dożycia przez pacjenta dnia 140.) wynosi 86,18%
* Niepowodzenie: **0,1382** – prawdopodobieństwo, że osoba wróci do nałogu przed początkiem badanego przedziału (prawdopodobieństwo śmierci pacjenta przed dniem 140) wynosi 13,82%
* Mediana dalszego trwania życia: **417,9** i błąd standardowy mediany: **27,36** – w badanym przedziale połowa pacjentów przeżyje czas nie dłuższy niż 417,9, a pozostała połowa czas dłuższy niż 417,9 dnia, z błędem na poziomie 27,36 dnia.
* Hazard: **0,0019** - ryzyko, że przypadek, który przeżył do początku danego przedziału ulegnie w tym przedziale zdarzeniu

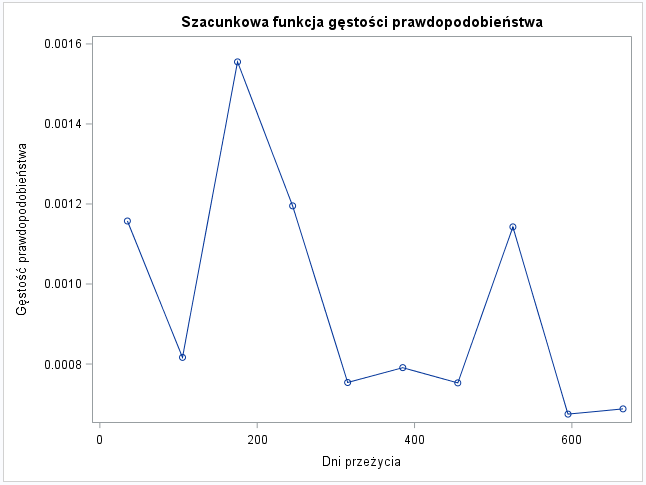


*Wykres 3. Krzywa przeżycia w zależności od zmiennej czasowej*

Krzywa spada monotonicznie w każdym przedziale. Na wykresie widać również przedział ufności 95%, który nie powiększa się znacznie wraz ze wzrostem zmiennej czasu. Nie występuje tu heteroskedastyczność, a co za tym idzie krzywa jest istotna statystycznie.



*Wykres 4. Funkcja hazardu dla metody tradycyjnej*



*Wykres 5. Funkcja gęstości prawdopodobieństwa dla modelu tradycyjnego*

Funkcja hazardu nie wykazuje się monotonicznością- ani nie rośnie ani nie maleje przez cały okres trwania badania. Występują tu 2 maksima- jedno przy ok. 180 dniach i drugie przy ok. 500 dniach. Biorąc pod uwagę rozkład zmiennej dni przeżycia z poprzedniego rozdziału, warto dodać, że ok. dnia 500. występuje znaczący wzrost danych ocenzorowanych, co ma znaczący wpływ na występujący w tym okresie odsetek niepowodzeń. Gdyby jednak przeprowadzić linię klasycznej regresji liniowej przez tę funkcję, należałoby stwierdzić, że funkcja hazardu rośnie w czasie, co oznacza że ryzyko porażki wzrasta w czasie.

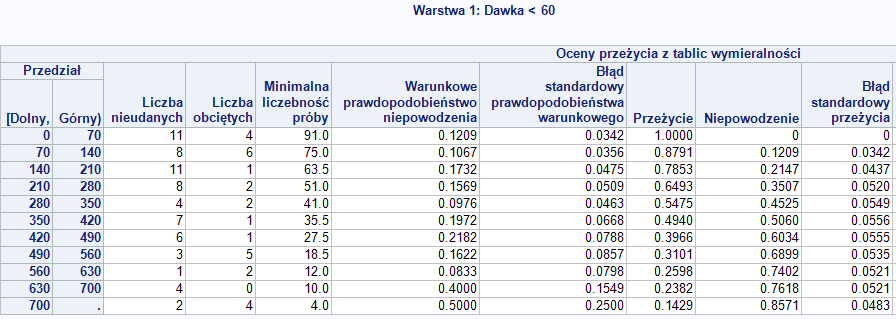
## 3.2. Metoda tradycyjna ze stratyfikacją

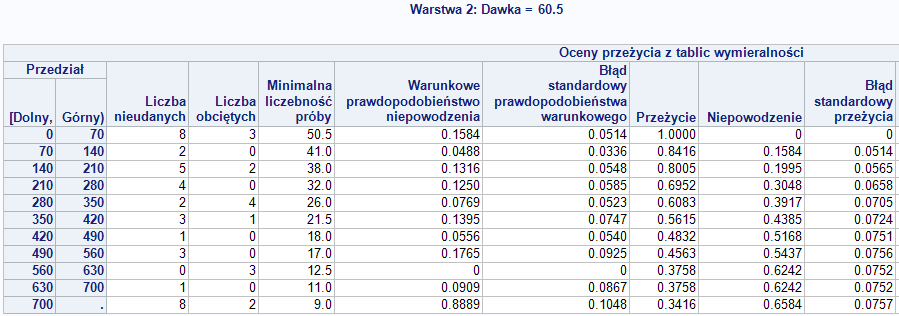
Kolejnym krokiem w analizie metodą tradycyjną jest wykorzystanie zmiennej warstwującej. Autor zdecydował się użyć jako zmiennej stratyfikującej zmiennej DAWKA, jako że patrząc po jej rozkładzie(patrz- wykres 2)- może ona mieć największy wpływ na zmienną STATUS. Poza tym, już po samej analizie danych zagregowanych można ocenić, że pobyt w więzieniu miał mały wpływ na powrót do nałogu, a pacjenci z kliniki 1 o wiele częściej doświadczali porażki niż pacjenci z kliniki 2.

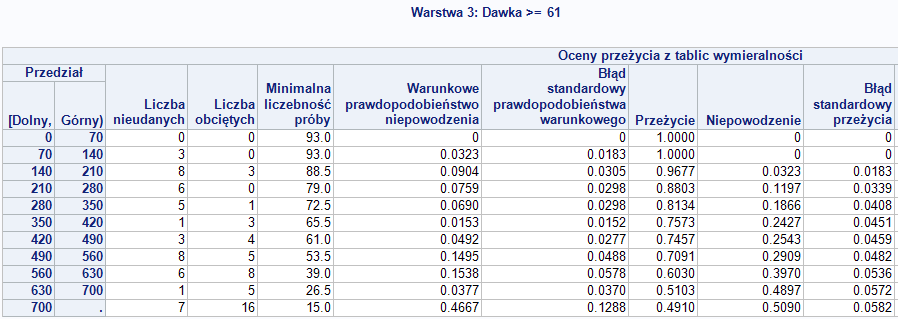
Zmienna została podzielona na 3 warstwy:

* wartości poniżej 60 mg/dzień
* wartość dominująca- 60 mg/dzień
* wartości powyżej 60 mg/dzień

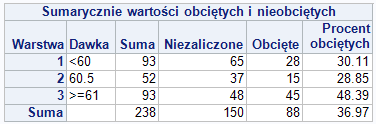
UWAGA: proszę nie sugerować się wartością 60.5 przy Warstwie 2- jest to umowny środek przedziału oszacowany przez SAS Enterprise Miner, ale uwzględnia tylko dawkę 60 mg/dzień.



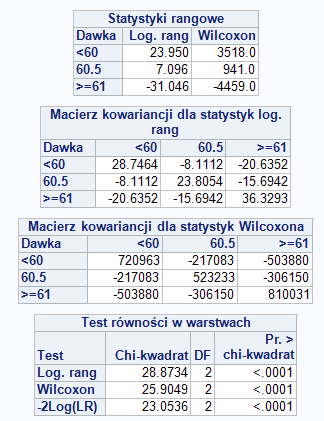




*Tabela 10. Oceny przeżycia dla warstw*



*Tabela 11. Sumaryczne wartości obciętych i nieobciętych*

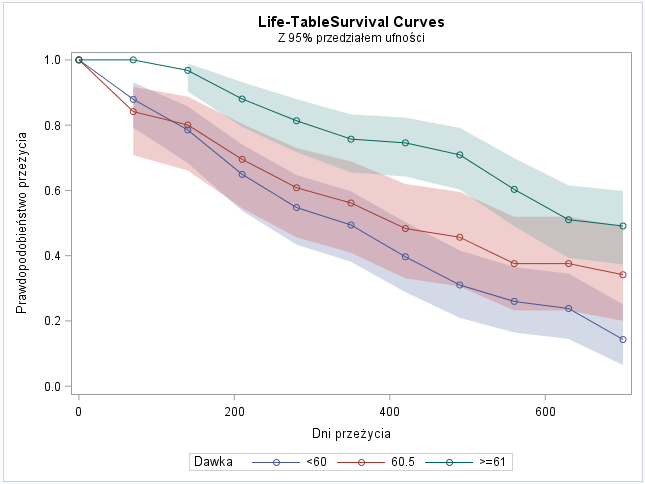


*Tabela 12. Testy homogeniczności krzywych przeżycia w warstwach*

Ponieważ funkcja przeżycia daje pełne rozeznanie doświadczenia przeżycia każdej grupy, naturalnym podejściem do odpowiedzi na to pytanie o homogeniczność jest przetestowanie hipotezy zerowej, stanowiącej że funkcje, które przeżyły są takie same w obu grupach; to jest, S1(t) = S2(t) dla wszystkich t, gdzie indeksy dolne rozróżniają dwie grupy. PROC LIFETEST oblicza dwie alternatywne statystyki do testowania tej hipotezy zerowej: test log-rank (znany również jako test Mantela-Haenszela) i test Wilcoxona. Trzeci test, statystyka ilorazu wiarygodności, jest obliczany przy dodatkowym założeniu, że czasy zdarzeń mają rozkład wykładniczy. Postawione hipotezy są w tym przypadku następujące:

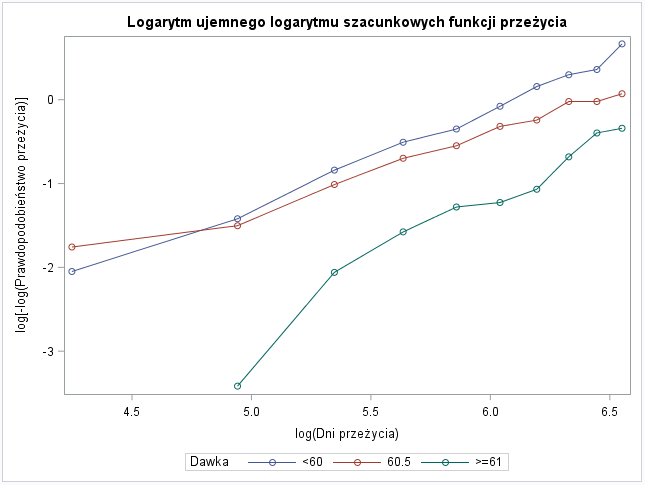
* H0: warstwy są homogeniczne
* H1: warstwy są heterogeniczne

Odrzucamy H0; funkcje dożycia w warstwach są heterogeniczne; istotnie statystycznie różnicuje się czas dożycia, czyli czas powrotu do nałogu.



*Wykres 6. Funkcja dożycia z podziałem na warstwy*

Powyższy wykres potwierdza wniosek z testów homogeniczności, mianowicie- warstwy są heterogeniczne. Na powyższym wykresie widać, że funkcja dożycia maleje najwolniej dla największych dawek (> 60 mg/dzień), a najszybciej dla dawek najmniejszych. Na tej podstawie można wysunąć wniosek, że im większa dawka metadonu, tym mniejsze prawdopodobieństwo powrotu przez pacjenta do nałogu.



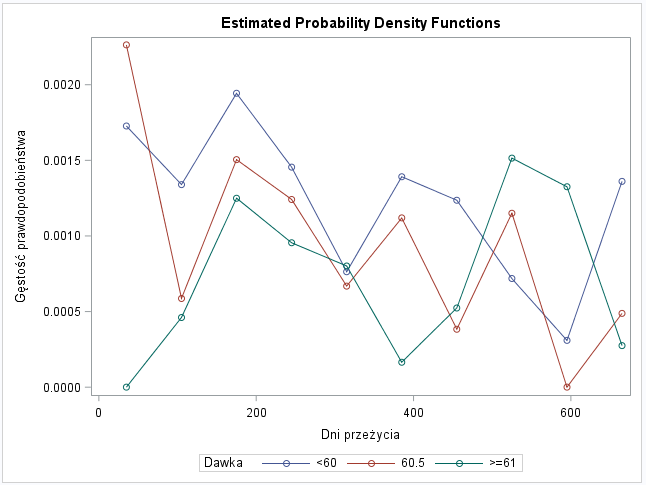
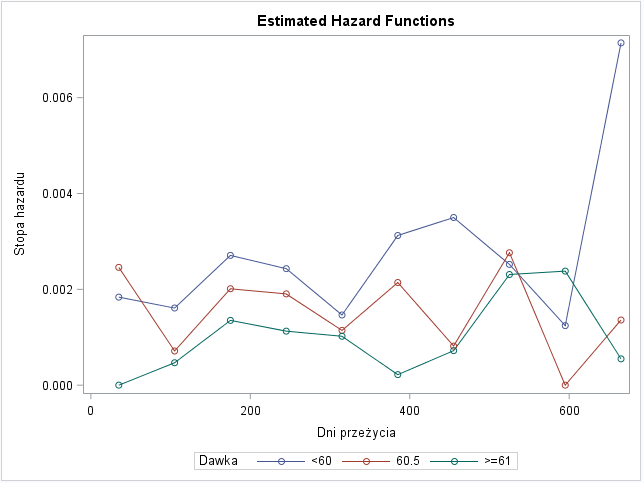
*Wykres 7. Log ujemnego logu funkcji przeżycia*

Rysunek pokazuje logarytm ujemnego logarytmu funkcji przeżycia. Aby sprawdzić w przybliżeniu, jak kształtuje się rozkład zmiennej czasowej, można posłużyć się poniższymi założeniami:

• Jeśli rozkład wykładniczy byłby odpowiedni, krzywa LLS powinna być linią prostą o nachyleniu równym 1;

• Jeśli zaś rozkład Weibulla byłby odpowiedni, krzywa LLS powinna być linią prostą o nachyleniu mniejszym bądź większym od 1.

W tym badaniu wykres LLS wygląda w przybliżeniu liniowo, jednak o nachyleniu mniejszym niż 1, co sugeruje graficznie, że rozważany rozkład czasu przeżycia to rozkład Weibulla – jednak na tym etapie analizy (analiza nieparametryczna) nie jesteśmy w stanie tego stwierdzić. Ze względu na wahania wykresu w końcowej fazie, być może konieczne będzie zastosowanie innego rozkładu, np. log-logistycznego lub log-normalnego.



*Wykres 8. Funkcje prawdopodobieństwa i hazardu dla metody tradycyjnej ze stratyfikacją*

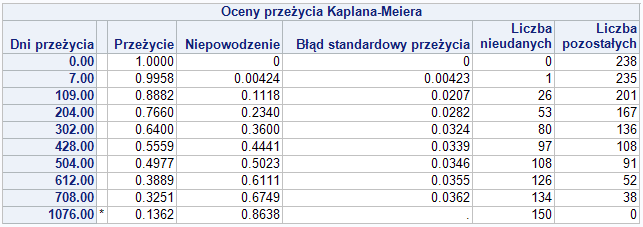
Obie funkcje cechują się brakiem monotoniczności. Można jednak zauważyć pewne podobieństwa: o okresie od 100 do 300 dni, obie funkcje dla wszystkich warstw zachowują się w podobny sposób. Może to oznaczać, że w tym okresie czynniki inne niż dawka mają wpływ na proces leczenia pacjenta. Po ok. 330 dniach funkcja prawdopodobieństwa ,,rozchodzi się” osiągając minimum globalne dla warstwy >60 mg/dzień. Następnie jednak, dla tejże warstwy funkcja szybko rośnie, osiągając swe maksimum około dnia 520, podczas gdy dla najmniejszej dawki funkcja znacząco maleje. Sugerować to może, że większe dawki mają sens tylko do około roku terapii, następnie najlepsze wydaje się zmniejszenie dawki. Wartości funkcji powyżej 600 dni są trudno interpretowalne przez małą liczebność próby w tym przedziale czasu.

## 3.3. Metoda Kaplana-Meiera

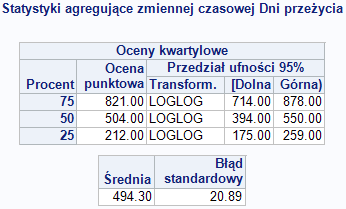
Kolejną metodą nieparametryczną analizy historii zdarzeń jest metoda Kaplana-Meiera. W odróżnieniu od metody tradycyjnej metoda ta nie wymaga arbitralnego przyjęcia punktów w czasie, w którym dokonywana jest estymacja funkcji przeżycia. Funkcja przeżycia liczona jest w każdym punkcje, w którym wystąpiło co najmniej jedno zdarzenie. Funkcja przeżycia estymowana jest w sposób następujący:

gdzie nj – liczba pacjentów, którzy nie wrócili do nałogu do momentu t­­­j­, dj­ – liczba pacjantów, którzy powrócili do nałogu w momencie tj.

Analiza rozpoczęta została od przeprowadzenia estymacji dla wszystkich 238 pacjentów w próbie. Z uwagi na dużą liczbę wydarzeń, tabela z wartościami funkcji przeżycia zawiera dużą liczbę informacji. Z tego powodu poniżej przedstawiono wybrane wartości tejże funkcji.

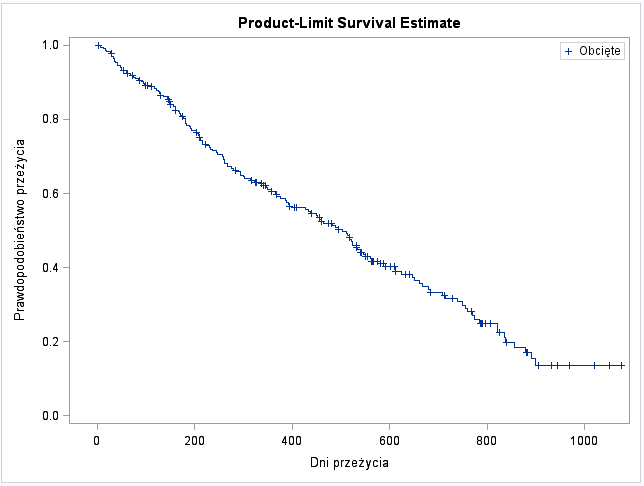


*Tabela 13. Wartości funkcji przeżycia dla metody Kaplana-Meiera*



*Tabela 14. Statystyki agregujące dla metody Kaplana-Meiera*

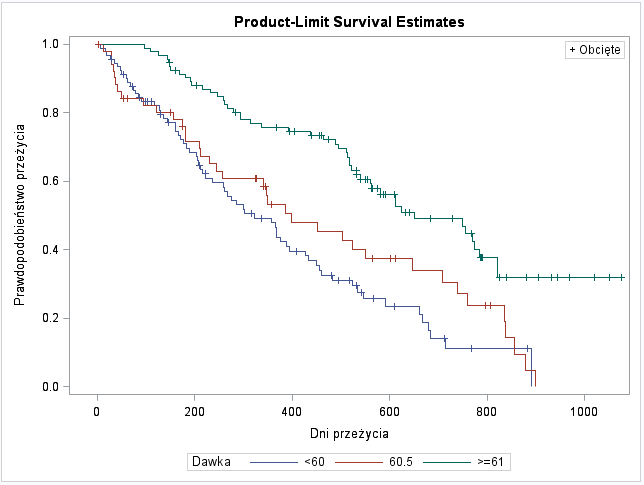
Średni czas dożycia wynosi 494 dni, średnia wynosi 504 dni. Średnia jest większa niż mediana, więc mamy do czynienia z rozkładem prawostronnie asymetrycznym.



*Wykres 9. Funkcja dożycia w metodzie Kaplana-Meiera*

Na podstawie wyestymowanej funkcji dożycia można stwierdzić, że prawdopodobieństwo dożycia maleje w tempie 10% na 100 dni.

Kolejnym etapem było użycie zmiennej stratyfikującej DAWKA- analogicznie do metody tradycyjnej. Wyniki testów homogeniczności okazały się być bliźniacze jak dla metody tradycyjnej, więc autor uznał ponowne publikowanie tych samych wyników za niepotrzebne.



*Wykres 10. Funkcja dożycia metodą Kaplana-Meiera ze zmienną stratyfikującą*

Interpretacja wyników przy użyciu zmiennej warstwującej DAWKA jest analogiczna jak przy metodzie tradycyjnej- funkcja dożycia maleje w mniejszym tempie dla wyższych dawek metadonu, a maleje szybciej dla dawek mniejszych. Nie występuje homogeniczność funkcji przeżycia w warstwach.

# 4. Porównanie modeli i wnioski

Próbka, na której budowane były analizy miała 238 obserwacji. Tak duża próba wystarczyła, aby dokonać estymacji ze stosunkowo dużą precyzją. W obu modelach błąd standardowy funkcji przeżycia dla górnych wartości zmiennej czasowej nie przekraczał wartości 8%. Świadczy to o dużej precyzji dokonanych szacunków.

Funkcja przeżycia w obu modelach miała zbliżony przebieg.

Głównym wnioskiem odkrytym dzięki stratyfikacji jest wpływ dawki metadonu na ryzyko powrotu do nałogu narkotykowego. Analiza pokazała, że większe dawki metadonu niż 60 mg/dzień korzystnie wpływają na terapię pacjenta objętego leczeniem uzależnienia od heroiny. Jednakże w dłuższym okresie czasu, najkorzystniejszą strategią jest stopniowe zmniejszanie dawki leku.

# 5. Bibliografia

1. SAS Institute; ” SAS/STAT 13.2 User’s Guide”; 2014

2. E.Frątczak, U.Sienkiewicz, H.Babiker; „Analiza historii zdarzeń – elementy teorii, wybrane przykłady zastosowań”; Szkołą Główna Handlowa w Warszawie; Wydanie 1 2009

3. https://www.ncbi.nlm.nih.gov/pubmed/1988793

# 6. Kody SAS

/\* Importowanie zbioru ADDICTS \*/

**DATA** WORK.ADDICTS;

LENGTH

Klinika **3**

Status **3**

'Dni przeżycia'n **8**

'Więzienie'n **3**

Dawka **3** ;

DROP

ID ;

LABEL

Klinika = " "

Status = " "

'Dni przeżycia'n = " "

'Więzienie'n = " "

Dawka = " " ;

FORMAT

Klinika BEST12.

Status BEST12.

'Dni przeżycia'n BEST12.

'Więzienie'n BEST12.

Dawka BEST12. ;

INFORMAT

Klinika BEST12.

Status BEST12.

'Dni przeżycia'n BEST12.

'Więzienie'n BEST12.

Dawka BEST12. ;

INFILE 'C:\Users\Sebastian\AppData\Local\Temp\SEG3012\ADDICTS-3d9e43b25aa64f6c82fe4d186a99dcc2.txt'

LRECL=**17**

ENCODING="WLATIN2"

TERMSTR=CRLF

DLM='7F'x

MISSOVER

DSD ;

INPUT

ID : $1.

Klinika : BEST32.

Status : BEST32.

'Dni przeżycia'n : BEST32.

'Więzienie'n : BEST32.

Dawka : BEST32. ;

**RUN**;

/\* Tabela jednoczynnikowa dla STATUS \*/

**PROC** **SQL**;

CREATE VIEW WORK.SORT AS

SELECT T.Status

FROM WORK.ADDICTS as T

;

**QUIT**;

**PROC** **FREQ** DATA=WORK.SORT

ORDER=INTERNAL

;

TABLES Status / SCORES=TABLE;

**RUN**;

/\* Statystyki agregujące dla STATUS i Dni przeżycia\*/

**PROC** **SQL**;

CREATE VIEW WORK.SORTTempTableSorted AS

SELECT T."Dni przeżycia"n, T.Status

FROM WORK.ADDICTS as T

;

**QUIT**;

**PROC** **MEANS** DATA=WORK.SORTTempTableSorted

FW=**12**

PRINTALLTYPES

CHARTYPE

QMETHOD=OS

NWAY

VARDEF=DF

MEAN

STD

MIN

MAX

N

Q1

MEDIAN

Q3 ;

VAR "Dni przeżycia"n;

CLASS Status / ORDER=UNFORMATTED ASCENDING;

**RUN**;

ODS GRAPHICS ON;

TITLE;

/\*-----------------------------------------------------

\* Generuje histogramy za pomocą PROC UNIVARIATE..

\*/

TITLE;

TITLE1 "Statystyki agregujące";

TITLE2 "Histogramy";

**PROC** **UNIVARIATE** DATA=WORK.SORTTempTableSorted NOPRINT ;

VAR "Dni przeżycia"n;

CLASS Status ;

HISTOGRAM ;

**RUN**; **QUIT**;

ODS GRAPHICS OFF;

/\* Tabela jednoczynnikowa dla KLINIKA \*/

**PROC** **SQL**;

CREATE VIEW WORK.SORTk AS

SELECT T.Klinika

FROM WORK.ADDICTS as T

;

**QUIT**;

**PROC** **FREQ** DATA=WORK.SORTk

ORDER=INTERNAL

;

TABLES Klinika / SCORES=TABLE;

**RUN**;

/\* Tabela zagregowane dla KLINIKA I STATUS \*/

**PROC** **TABULATE**

DATA=WORK.ADDICTS

;

CLASS Klinika / ORDER=UNFORMATTED MISSING;

CLASS Status / ORDER=UNFORMATTED MISSING;

TABLE /\* Wymiar wiersza \*/

Klinika,

/\* Wymiar kolumny \*/

Status\*

N

N ;

;

**RUN**;

/\* Zagregowana tabela dla zmiennych więzienie i status \*/

**PROC** **TABULATE**

DATA=WORK.ADDICTS

;

CLASS Status / ORDER=UNFORMATTED MISSING;

CLASS "Więzienie"n / ORDER=UNFORMATTED MISSING;

TABLE /\* Wymiar wiersza \*/

'Więzienie'n\*

N,

/\* Wymiar kolumny \*/

Status ;

;

**RUN**;

/\* Rozkład zmiennej DAWKA z uwzględnieniem zmiennej STATUS \*/

**PROC** **SQL**;

CREATE VIEW WORK.SORTTempTableSorted AS

SELECT T.Dawka, T.Status

FROM WORK.ADDICTS as T

;

**QUIT**;

Legend1

FRAME

;

Axis1

STYLE=**1**

WIDTH=**1**

MINOR=NONE

;

Axis2

STYLE=**1**

WIDTH=**1**

;

TITLE;

TITLE1 "Wykres słupkowy";

**PROC** **GCHART** DATA=WORK.SORTTempTableSorted

;

VBAR

Dawka

/

SUBGROUP=Status

CLIPREF

FRAME TYPE=FREQ

LEGEND=LEGEND1

COUTLINE=BLACK

RAXIS=AXIS1

MAXIS=AXIS2

;

**RUN**; **QUIT**;

TITLE;

/\* Współczynniki korelacji Pearsona \*/

**PROC** **SQL**;

CREATE VIEW WORK.SORTTempTableCORR AS

SELECT T.Klinika, T."Więzienie"n, T.Dawka

FROM WORK.ADDICTS as T

;

**QUIT**;

TITLE1 "Analiza korelacji";

**PROC** **CORR** DATA=WORK.SORTTempTableCORR

PLOTS=NONE

PEARSON

VARDEF=DF

;

VAR Klinika "Więzienie"n Dawka;

**RUN**;

TITLE;

/\* Tabele przeżycia metodą tradycyjną \*/

ODS GRAPHICS ON;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

**PROC** **SQL**;

CREATE VIEW WORK.TMP0TempTableInput AS

SELECT T."Dni przeżycia"n, T.Status, T.Dawka, T."Więzienie"n

FROM WORK.ADDICTS as T

;

**QUIT**;

/\*-----------------------------------------------------

Sekcja analizy

\*/

TITLE;

TITLE1 "Analiza tabel przeżycia";

/\*-----------------------------------------------------

Uruchom PROC LIFETEST, aby przeprowadzić analizę.

\*/

**PROC** **LIFETEST** DATA=WORK.TMP0TempTableInput

ALPHA=**0.05**

METHOD=LIFE

INTERVALS= **70** **140** **210** **280** **350** **420** **490** **560** **630** **700** PLOTS(ONLY)=SURVIVAL( CL )

PLOTS(ONLY)=LOGSURV

PLOTS(ONLY)=LOGLOGS

PLOTS(ONLY)=HAZARD

PLOTS(ONLY)=DENSITY

;

TEST "Więzienie"n;

TIME "Dni przeżycia"n \* Status (**0**);

**RUN**;TITLE;

**RUN**; **QUIT**;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

TITLE;

ODS GRAPHICS OFF;

/\* Tabele przeżycia metodą tradycyjną ze stratyfikacją \*/

ODS GRAPHICS ON;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

**PROC** **SQL**;

CREATE VIEW WORK.TMP0TempTableInput AS

SELECT T."Dni przeżycia"n, T.Status, T.Dawka, T."Więzienie"n

FROM WORK.ADDICTS as T

;

**QUIT**;

TITLE;

TITLE1 "Analiza tabel przeżycia";

/\*-----------------------------------------------------

Uruchom PROC LIFETEST, aby przeprowadzić analizę.

\*/

**PROC** **LIFETEST** DATA=WORK.TMP0TempTableInput

ALPHA=**0.05**

METHOD=LIFE

INTERVALS= **70** **140** **210** **280** **350** **420** **490** **560** **630** **700** PLOTS(ONLY)=SURVIVAL( CL )

PLOTS(ONLY)=LOGSURV

PLOTS(ONLY)=LOGLOGS

PLOTS(ONLY)=HAZARD

PLOTS(ONLY)=DENSITY

;

STRATA Dawka ( **60** **61**);

TEST "Więzienie"n;

TIME "Dni przeżycia"n \* Status (**0**);

**RUN**;TITLE;

**RUN**; **QUIT**;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

TITLE;

ODS GRAPHICS OFF;

/\* Metoda Kaplana-Meiera \*/

ODS GRAPHICS ON;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

**PROC** **SQL**;

CREATE VIEW WORK.TMP0TempTableInput AS

SELECT T."Dni przeżycia"n, T.Status, T.Dawka

FROM WORK.ADDICTS as T;

**QUIT**;

TITLE;

TITLE1 "Analiza tabel przeżycia";

**PROC** **LIFETEST** DATA=WORK.TMP0TempTableInput

ALPHA=**0.05**

INTERVALS= **100** **200** **300** **400** **500** **600** **700** PLOTS(ONLY)=SURVIVAL

PLOTS(ONLY)=LOGLOGS;

TIME "Dni przeżycia"n \* Status (**0**);

**RUN**;TITLE;

**QUIT**;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

ODS GRAPHICS OFF;

/\* Metoda Kaplana-Meiera ze stratyfikacją\*/

ODS GRAPHICS ON;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

**PROC** **SQL**;

CREATE VIEW WORK.TMP0TempTableInput AS

SELECT T."Dni przeżycia"n, T.Status, T.Dawka

FROM WORK.ADDICTS as T;

**QUIT**;

TITLE;

TITLE1 "Analiza tabel przeżycia";

**PROC** **LIFETEST** DATA=WORK.TMP0TempTableInput

ALPHA=**0.05**

INTERVALS= **100** **200** **300** **400** **500** **600** **700** PLOTS(ONLY)=SURVIVAL

PLOTS(ONLY)=LOGLOGS;

STRATA Dawka ( **60** **61**);

TIME "Dni przeżycia"n \* Status (**0**);

**RUN**;TITLE;

**QUIT**;

%***\_eg\_conditional\_dropds***(WORK.TMP0TempTableInput,

WORK.TMP2TempTablePlot0);

ODS GRAPHICS OFF;