1.Ugolniony kappa, kappa,

2.tabela z wzorami,

3. kodowanie efekt,

3. kto odkryl funkcje logistyczna na nowo,

Or w modelu z interakcjami

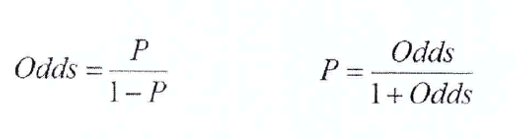
zapisac model uporzadkowany

Pytania z kolokwium semestr letni 2017/18 (z grupy FB)

1. zadanie 1 na uzupełniane wzorów i właściwości

O które wzory może chodzić?

Tabelka z

1. zadanie 2 z zależnością odds i P
2. zadanie 6 tyle że model z dodatkową jedną interakcją
3. podany model i na jego podstawie obliczyć punktowe oszacowanie OR i przedziały ufności, i chyba przy tym zadaniu trzeba było napisać czy jest to metoda dokładna czy surowa

CzY WIE KTOŚ O CO CHODZI Z surową i dokładną ? Gdzie o tym jest ?

1. napisać jakie statystyki służą do weryfikacji istotności

Str. 60 podrecznik (w skrypcie jest więcej, just sayin’) skrypt? Jest coś prócz podręcznika i slajdów? miałam na myśli slajdy ;p

-test wskaźnika wiarygodności

-test Walda

-test punktowy

-estymacja przedziałowa:

-dla jednego współczynnika

-dla interakcji

// w podręczniku są jeszcze wspomniane: test ilorazu wiarygodności (inna nazwa na test wskaźnika wiarygodności) i test punktowy

1. podana tablica kontyngencji i napisać wzór na odds i RR



OR = ad/bc

RR ==

1. napisać model logistyczny w postaci wielomianu i dlaczego jest albo nie jest spoko <coś takiego, nie pamiętam>

 to? A to nie chdzi o model wielomianowy? Bo to jest EVW :P Czyli tylko to pierwsze? Czy jeszcn ze coś innego? Hm wydaje mi się, że chodzi o ten model, który był po uporządkowanym? Niech ktoś potwierdzi :-)o

Niech ktoś napisze co tu powinno być.

A nie ? I tworzymy takich modeli G-1.

g i - indeksy dolne ofc

Gdzie:

g- badana kategoria

G- ilość kategorii

Wg mnie w przypadku zmiennej wynikowej uporządkowanej nie jest spoko, ponieważ tracimy informacje na temat tego porządku, ale dla zmiennej nominalnej chyba jest ok.

Niech ktoś potwierdzi/ zaprzeczy.

Mi się wydaje ze tu chodzi o to ze jest tabelka ze zmiennymi i trzeba napisac rownanie regresji liniowej, tak jak w tym zadaniu 6 podpunkt 1. Bo tamte postacie to są logitowe a nie wielomianowe. NA STRONIE 57 W PODRĘCZNIKU JEST TO ZAPISANE. SZACUJEMY G-1 MODELI.

1. co to jest lambda i jakie są właściwości asymptotyczne i symetryczne <nie jestem pewna jak brzmiało całe polecenie>

(*z podręcznika str 37*) Asymetryczny współczynnik lambda interpretuje się jako prawdopodobne polepszenie prognozy B dzięki znajomości A. Oblicza się go następującym wzorem:

Im bliższa jedności jest wartość współczynnika, tym silniejsza zależność B od A. Symetryczny współczynnik lambda otrzymuje się nijako uśredniając z ƛ(B|A) i ƛ(A|B).

Jest to miernik wzajemnej zależności badanych cech.

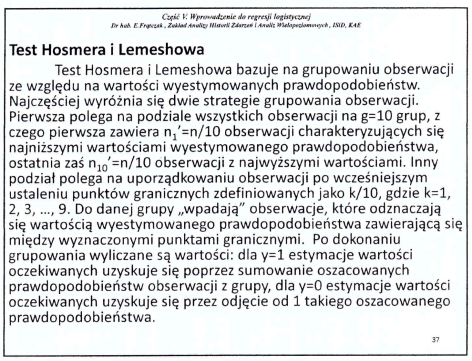
1. kto stworzył probit a kto logit

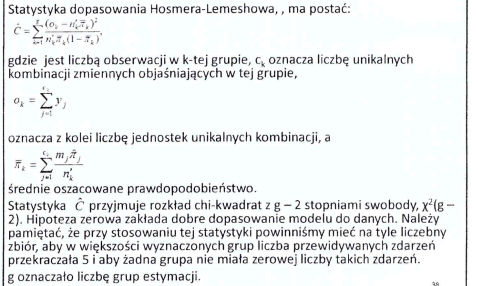
Probit - Bliss(1934) pomysł i nazwa “probit”, Finney - jakoś rozwinął.

Logit - Joseph Berkson - model logitowy 1944

1. opisać Hosmera-Lemeshowa

**Test Hosmera-Lemeshowa –** Test weryfikuje hipotezę o równości wartości obserwowanych i przewidywanych. Jeśli wartości obserwowane i przewidywane są wystarczająco bliskie, wówczas można założyć, że model jest dobrze dopasowany do danych. *Brak istotności tego testu jest pożądany. Jest to główna statystyka diagnostyczna modelu regresji logistycznej. Test Hosmer Lemeshow jest niezbędną statystyką podczas oceny dopasowania modelu do przewidywanych danych.* Na podstawie testu Hosmera Lemeshowa rozpatruje się rewizję modelu, jego przekształcenie lub zastosowanie alternatywnej metody analizy.





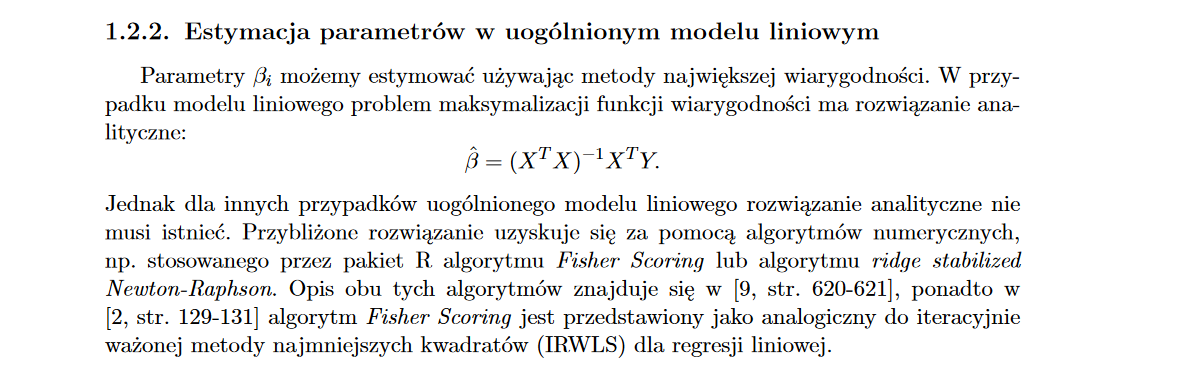
1. czym można estymować model regresji logistycznej i regresji liniowej i opisać te metody

Model regresji logistycznej jest szacowany przy WYKORZYSTANIU Metody Największej Wiarygodności, a regresji liniowej przy jej ZASTOSOWANIU.metody największyc kwadrató

Pytanie: Z tego co mi się wydaje, to do regresji liniowej najczęściej wykorzystuje się estymator MNK, ale nie widziałam nic o tym w skrypcie. JEST TO GDZIEŚ?! Czy po prostu to jest oczywiste i mamy to napisać?

Str. 56 w podręczniku -> model logitowy można szacować MNW

Z jakiejś strony internetowej:



1. była podana tablica jak na slajdach ze zmiennymi sztucznymi i uzupełnić luki (w środku puste) dla kodowania referencyjnego i napisać jak się interpretuje
2. napisać jak się liczy współczynnik Gamma, Kendalla i Somersa i jaka interpretacja

Statystyka Gamma: (nc - nd) / (nc+nd)

Statystyka D-Sommersa: (nc - nd) / t

Statystyka Tau Kendalla: (nc - nd) / (0,5\*n(n-1))

nc - liczba par zgodnych

nd - liczba par niezgodnych

t - liczba wszystkich par (bez tied -> czyli 00 i 11)

n - liczba obserwacji

Interpretacja: Statystyki przyjmują wartości od -1 do 1; wartości bezwzględne statystyk bliskie 1 wskazują na bardzo silne powiązanie zmiennych

1. Kto ponownie wynalazł funkcję logistyczną w 1920 r

Pearl i Reed,

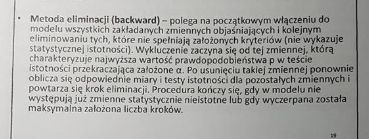
Egzamin 2017/2018

- definicja uogolnionych modeli liniowych - czy regresja logistyczna się do nich zalicza

- liczenie OR i RR z tablicy kontyngencji. Pytanie czy to była ocena ilorazu surowa czy dopasowana

- policzyć iloraz szans dopasowany i przedziały ufności w modelu z interakcja 2 stopnia, podane betaX1=-0.4, blad std=0.15 i betaX2=0.03 błąd std=0.012. X2 zmienna ciągła. X1 zmiana z 1 na 0 (lub odwrotnie) a X2=10.

- opisać selekcję backward



- opisać statystyki reszt wiarygodności i wartości wpływowych

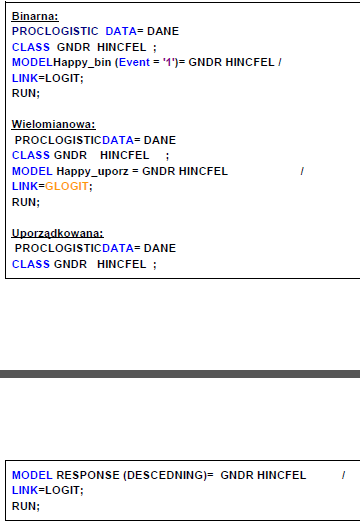
- opisać AIC i SC - do czego służą i jak są skonstruowane

Praktyka

- kod binarna

- kod uporządkowana

- kod wielomianowa



- wady wielomianowej wzgledem uporzadkowanej

- wyliczać expB, oceniać czy istotne czy nie istotne na podstawie przedziałów ufności i pvalue

- opisać i zinterpretować czułość i specyficznosc na przykładowych danych (tablica klasyfikacji - jakim kodem się ją dodaje w sasie)

- opisać czym się wybiera punkt odcięcia- krzywa ROC i jak ona jest zbudowana

- interpretacja wyników każdej z trzech regresji

- interpretacja wyników testu proporcjonalnosci

- do czego służy Hosmer Lemeshow i jak to włączyć w SASie w model

LACKFIT

Opiera sie na statystyce chi-kwadrat, jeśli piValue > poziom istotność(0,05) to model dobrze dopasowany do danych. Testowana była hipoteza zerowa o tym że model dobrze dopasowany do danych.

- jaka opcja włączyć liczenie przedziałów ufności walda i jak zmienic jednostkę zmiennej