Neil Yetz

02/11/2016

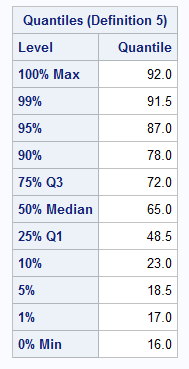
ERHS 642: Applied Logistic Regression

Dr. Bachand

**ERHS 642 Logistic Regression Spring 2016 Homework Assignment 2 – New Version**

1. In a temporary ICU data set, categorize the AGE variable into 4 categories based on the quartiles of AGE and create design variables for the categorized AGE variable. Then,

Table 1.1: Quantiles of AGE Variable



Age categories:

1 = 16-48.4

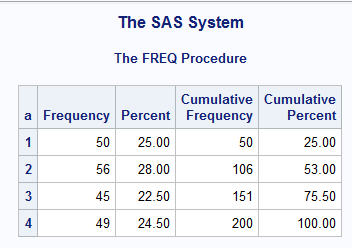
2 = 48.5-64.9

3= 65-71.9

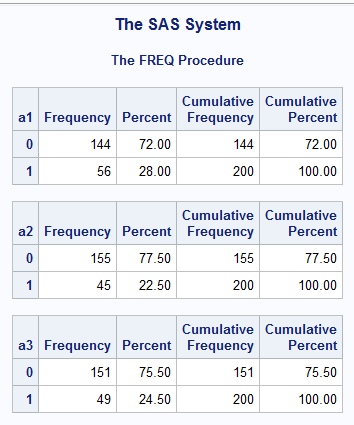
4 = 72-92

1. Create a frequency table for the 4-category AGE variable.

Table 1.2: Frequency table for 4-category age (a) variables

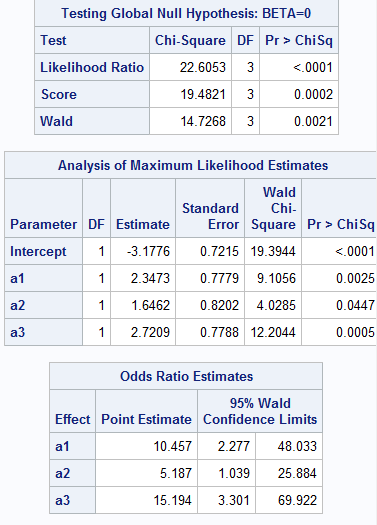


1. Create a frequency table for each design variable and compare the results to the results in 1a. How do the frequency tables for the design variables relate to the frequency table for the 4-category AGE variable?

Table 1.3: Frequency table for age (a) vs all other categories.

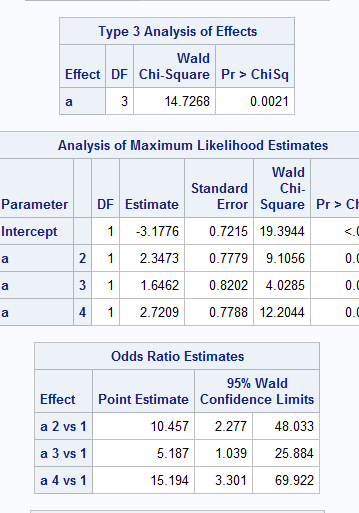
1. Perform logistic regression with outcome variable STA and the design variables you created as the independent variables.

Table 1.4: Results from proc logistic with self-created design variables



1. Perform logistic regression with outcome variable STA and the 4-category AGE variable as the independent variable using the class statement (i.e. let SAS create the design variables). Compare the results to the results in

Table 1.5: Results from proc logistic with SAS-created design variables

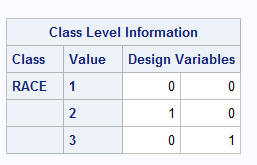


**Results in tables 1.3 and 1.4 are identical.**

1c. 2. H-L chapter 2, page 46-47, exercise 2

(a)

Table 2.1: Design variables for race



(b)

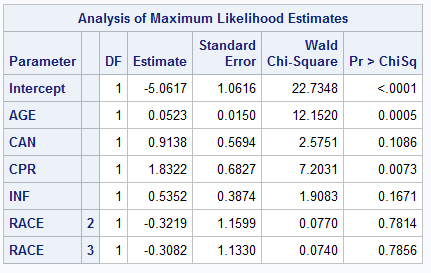
, where x = (age, can, cpr, inf, R2, R3)

, where x = (age, can, cpr, inf, R2, R3)

**There are 7 parameters within this model including the intercept.**

(d)

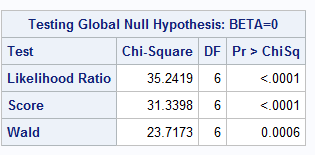
Table 2.2: Analysis of Maximum Likelihood Estimates



, where x = (age, can, cpr, inf, R2, R3)

(e) Compare model with all variables (age, can, cpr, inf and race) to model with no variables. You don’t have to answer the question about assumptions or the deviance

Table 2.3: Likelihood Ratio, Score Test, & Wald Test for Full model



Likelihood Ratio Test: G = 35.2419; DF = 6; p=<.0001

(f)

Table 2.4: Analysis of Maximum Likelihood estimates for reduced model: CAN, INF, Race 2 & Race 3 removed.

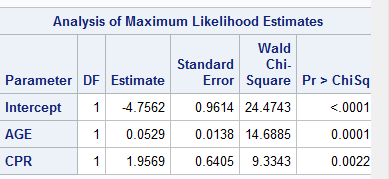


Table 2.5: Likelihood Ratio, Score Test, & Wald Test for Full model

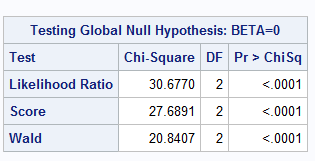
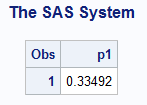


Table 2.6: proc print of χ2 p-value at 4-degrees of freedom for value 4.5649.

****

G (full model vs reduced model) = 35.2419 – 30.6770 = **4.5649**; df = **4**;p = **0.33492**

Don’t do parts c, g, h

libname sdat 'C:\Users\ndyet\_000\Desktop\Class Folders\Spring 2016\ERHS 642\Data';

**data** sdat.icu\_altered; set icu\_altered; **run**;

**data** icu\_altered; set sdat.icu\_altered;

\* Question 1;

if **0**<age<**48.5** then a=**1**;

else if **48.5**<=age<=**65** then a=**2**;

else if **65**<age<=**72** then a=**3**;

else if **72**<=age<=**92** then a=**4**;

if a=**1** then do; a1=**0**; a2=**0**; a3=**0**; end;

else if a=**2** then do; a1=**1**; a2=**0**; a3=**0**; end;

else if a=**3** then do; a1=**0**; a2=**1**; a3=**0**; end;

else if a=**4** then do; a1=**0**; a2=**0**; a3=**1**; end;

**run**;

**proc** **print** data=icu\_altered; **run**;

**proc** **univariate** data=icu\_altered; var age; **run**;

\* Question 1a;

**proc** **freq** data=icu\_altered; tables a; **run**;

\*Question 1b;

**proc** **freq** data=icu\_altered; tables a1 a2 a3; **run**;

\* 1.c;

**proc** **logistic** descending data=icu\_altered; model sta=a1 a2 a3; **run**;

\* 1.d;

**proc** **logistic** descending data=icu\_altered; class a/param=ref ref=first; model sta=a;

**run**;

\* Question 2;

\* 2-2.d, e, f;

**proc** **logistic** descending data=icu\_altered;

class race/param=ref ref=first;

model sta=age can cpr inf race;

**run**;

\* 2-2.f;

**proc** **logistic** descending data=icu\_altered;

model sta=age cpr;

**run**;

**data** pval;

p1=**1**-probchi(**4.5649**,**4**);

**run**;

**proc** **print** data=pval; **run**;