Homework 7

Aji John

Winter Quarter 2018

We are interested in assessing associations among mortality, creatinine, age, sex, race, and smoking behavior in a population of generally healthy elderly subjects in four U.S. communities

Question 1

Suppose we are interested in evaluating associations with all-cause mortality using a logistic regression model. Perform a logistic regression analysis with indicator of death within 5 years of study enrollment as the response and with creatinine, age, indicator for ever smoked, and indicator of white race as predictors.

(a)

Provide an interpretation of the exponentiated intercept term in the logistic regression model.

Intercept is the mean odds of dying in 5 years for newborn non-white race with CRT of zero who never smoked.

(b)

Provide an interpretation of the exponentiated age slope in the logistic regression model.

The odds ratio

(c)

From the logistic regression model, is there evidence of an association between death within 5 years of study enrollment and creatinine after adjusting for the other predictors? Give full inference.

(d)

From the logistic regression model, is there evidence of an association between death within 5 years of study enrollment and age after adjusting for the other predictors? Provide full inference.

(e)

From the logistic regression model, what is the best estimate of the odds of dying within 5 years of study enrollment for a non-white race individual who is 76 years old, has previously smoked, and has a creatinine level of 0.95.

(f)

From the logistic regression model, what is the best estimate of the probability of dying within 5 years of study enrollment for a white race individual who is 69 years old, has never smoked, and has a creatinine level of 1.2.

Question 2

Now perform a logistic regression analysis with indicator of death within 5 years of study enrollment as the response and creatinine, age, indicator of ever smoked, indicator of white race, and sex as predictors.

(a)

Provide an interpretation of the exponentiated intercept term in the logistic regression model.

It is the base odds for for group which does not die in 5 years for new-born females of non-smoking non-white race with CRT of 0. Not scientifically relevant as not the study group.

(b)

Provide an interpretation of the exponentiated sex slope in the logistic regression model.

Odds ratio of two groups dying in 5 years with not dying in 5 years between males and females. The odds ratio is for the two sex groups who have same race, smoking status, age and CRT level.

(c)

Provide full inference for an association between all-cause mortality within 5 years and sex using the logistic regression model.

(d)

Provide full inference for an association between all-cause mortality within 5 year years and creatinine using the logistic regression.

(e)

Is sex a confounder, precision variable, both or neither for the associations between all-cause mortality within 5 years of study enrollment and each of the other four predictors of interest: creatinine, age, indicator of ever smoked, and indicator of white race? Explain and provide evidence to support your reasoning.

Race looks to be a precision variable.

- As it reduces the variance when race is added, the confidence intervel becomes tighter(evident in scatter plot)

Question 3

Now suppose we are interested in evaluating associations with all-cause mortality using a Poisson regression analysis. Perform a Poisson regression analysis with indicator of death within 5 years as the response and creatinine, age, indicator of ever smoked, indicator of white race, and sex as predictors.

(a)

Provide an interpretation of the exponentiated intercept term of the Poisson regression model.

Exponentiated intercept(.00141) is the base rate of death in 5 years for newborn female non-white race who have never smoked, and have CRT as 0 mm/dl.

(b)

Provide an interpretation of the exponentiated creatinine slope in the Poisson regression model.

Exponentiate CRT slope (2.95) is the odds ratio or the risk-rate(RR) for subjects dying in 5 years/not dying in 5 year who differ by one unit change in CRT by keeping all the other covariates constant.

(c)

Provide full inference with the Poisson regression model for an association between all-cause mortality within 5 years of enrollment and creatinine after adjusting for the other predictors.

When comparing two groups(dying in 5 vs not dying in 5 years) with CRT differing in 1 unit(mg/dl), but having the same sex, age, racial group and smoking status, the risk of 5-yr all-cause mortality is estimated to be 111.4% lower in group which does not die in 5 years. Risk-Rate is 2.11 (95% CI 1.51 to 2.95 with p-value < 0.001)

(d)

Provide full inference with the Poisson regression model for an association between all-cause mortality within 5 years of enrollment and sex after adjusting for the other predictors.

When comparing two groups(dying in 5 vs not dying in 5 years) between males and females, but having the same CRT, age, racial group and smoking status, the risk of 5-yr all-cause mortality is estimated to be 35.2% lower in group which does not die in 5 years. Risk-Rate is 1.35 (95% C.94 to 1.93 with p-value = 0.09)

Not right, need analysis