HOMEWORK 4 & 5
Due 3/15/05

For this homework use the data set icu.dta, available on the course Web site. These data were collected on 200 patients who were part of a larger study investigating in-hospital survival of patients admitted to an ICU. A description of the study and variables is included on the next page. As in previous assignments, please cut and paste output and graphs into a Word document for printing and submission. Use `Courier` font to preserve formatting of any Stata output. Please hand in a hard copy to Olivia. The due date is March 15, 2005.

1. Use the `logistic` command to fit a logistic regression model relating the binary outcome variable `sta` (vital status) to `age` and `crn` (history of chronic renal failure). Assume that the relationship between the log odds of mortality (`sta`) and `age` is sufficiently linear so that nonlinearity is not a concern in modeling.

Answer the following questions about the output produced:
   A. State the interpretation of the reported odds ratios for `age` and `crn`.
   B. Interpret the z statistic and associated P-value for `crn`.
   C. Interpret the likelihood ratio chi-squared test provided (LR chi2) for the model.
   D. Perform two likelihood ratio tests, one each for `crn` and `age`. (These tests can be performed with the `lrtest` command as demonstrated in lecture and lab #10, and compare the model including both predictors to the reduced model excluding the variable of interest.) Compare the results to the Wald (z) tests for each variable from the joint model. Based on the improvement in likelihood associated with including each variable in the joint model, which variable appears to be the better predictor of vital status following admission to the ICU in this study?

2. Perform additional logistic analyses to assess whether or not `age` confounds the relationship between `sta` and `crn`.

3. Fit a logistic model for vital status (`sta`) including the variables `ser` and `age`, and allowing for an interaction between these two predictors. As in the first exercise, you may assume that a linear relationship between the log odds of the outcome and `age` is adequate in patients in both groups defined by the variable `ser`.

   A. Use the likelihood ratio test to evaluate the significance of the interaction. Compare to the results of the Wald test for the interaction term and comment on any differences.
   B. Calculate the estimated odds ratios and associated 95% confidence interval for a ten-year increase in `age` for individuals with `ser=1`. Repeat for individuals with `ser=0`.
   C. Plot the relationship between the log odds of the outcome and `age` in both groups defined by the variable `ser`, and comment on any differences.

icu.dta: 200 observations, 21 variables

Description:

The ICU data set consists of a sample of 200 subjects who were part of a much larger study on survival of patients following admission to an adult intensive care unit (ICU). The major goal of this study was to develop a logistic regression model to predict the probability of survival to hospital discharge of these patients and to study predictors associated with ICU mortality.

Description	Codes/Values	Variable name
Identification Code	ID number	id_Vital Status 0 = Lived, 1 = Died
	sta	
Age	Age in years	age
Gender	0 = Male, 1 = Female	sex
Ethnicity	1 = White	race
	2 = Black	
	3 = Other	
Service at admission	0 = Medical	ser
	1 = Surgical	
Cancer part of problem	0 = No, 1 = Yes	can
History of chronic	0 = No	crn
renal failure	1 = Yes	
Infection at admission	0 = No	inf
	1 = Yes	
CPR prior to admission	0 = No, 1 = Yes	cpr
Syst. blood pressure	mm Hg	sys
Heart rate	beats/min	hra
Previous admission to ICU	0 = No	pre
within 6 months	1 = Yes	
Type of admission	0 = Elective	typ
	1 = Emergency	
Major bone fracture	0 = No	fra
at admission	1 = Yes	
PO2 from initial blood	0 ≥ 60	po2
gases	1 < 60	
PH from initial blood	0 > 7.25	ph
gases	1 ≤ 7.25	
PCO2 from initial blood	0 > 45	pco
gases	1 ≤ 45	
Bicarbonate from initial	0 > 18	bic
blood gases	1 ≤ 18	
Creatinine from initial	0 < 2.0	cre
blood gases	1 ≥ 2.0	
Level of conciousness	0 = No coma/stupor	loc
	1 = Deep stupor	
	2 = Coma	

REFERENCES

Hosmer and Lemeshow, *Applied Logistic Regression*, Wiley, (1989)

Lemeshow, S., Teres, D., Avrunin, J. S., Pastides, H. (1988). Predicting the Outcome of Intensive Care Unit Patients. Journal of the American Statistical Association, **83**, 348-356.