Homework
Applied Logistic Regression

WEEK 8

Exercise 1:

Consider a multivariable model for the ICU study with the variables: **LOC2 TYP AGE CAN SYS** and **AGExSYS**. Assess the fit of the model.

Exercise 2:

Use the hyponatremia.dta dataset.

(a) Fit a regression model with **female**, **urinat3p**, **runtime**, **wtdiff** and **bmi**. Evaluate the fit of the model using both Hosmer-Lemeshow test (using deciles of risk) and the Pearson Chi-square statistic. Assess whether the results of the two tests are consistent.

(b) On the basis of the logistic model with **runtime**, **bmi**, **bmi2** and **wtdiff** as covariates, estimate the sensitivity and specificity of classifying subjects as having or not having hyponatremia using the cut-off values for the probability of hyponatremia of 0.5.

(c) Repeat the previous exercise using different cut-off values for the probability of hyponatremia. Draw by hand the ROC curve using these values of sensitivity and specificity.

(d) Use Stata to obtain the ROC curve. What is the discriminatory power of the model?

(e) Suppose someone had fraudulently access to your PC and altered data of the dependent variable in such a way that the coefficients of the model would remain the same. However, the predicted probabilities of the outcome would be largely affected. What would happen to goodness-of-fit statistics?

(f) Fit a model with **female** and **urinat3p** as covariates. Assess the overall fit of the model and its discriminatory power by computing the Pearson Chi square goodness of fit statistic and the area under the ROC curve.

(g) Estimate the predicted probability of the outcome.