

## Assignment 5: Selected solutions

December 12, 2021

- Problem 1 (see R file):

- (a) See R file. The model is significant. You could either check the  $R^2 = 0.8837$  or use the deviance do the hypothesis testing (compared with null model).
- (b) 95% CI for  $\beta_{\text{Cl.thickness}}$  :  $[0.2566559, 0.8133641]$  (or  $[-0.8133641, -0.2566559]$  if you encode benign as 1 and malignant as 0). z-test for  $\beta_{\text{Cell.shape}}$  would not reject  $H_0$ .
- (c) See R file. Note that you need to *compare* it to the full model. You can either use the G statistic to do the hypothesis test (the reduced model (c) is not adequate with  $\alpha = 0.05$  while it is adequate with  $\alpha = 0.01$ ) or AIC/BIC to do the comparison.
- (d) The probability of this patient's tumor being benign= 0.1506149. Be careful how you encode the class label would affect the result you get via directly using '*predict*' function in R.
- (e) See R file. The model in part (c) would give the lowest AIC, comparing to all its subset models, which is 124.566.