*Confounding example = Example* in the book = If they have THE SAME amount of balance (student/non-student), then THE STUDENT will have lesser chance to default. In contrast IN GENERAL if we look at the model of only student, the parameter is positive and it is because STUDENTS HAVE GREATER BALANCE THAN NON-STUDENTS IN GENERAL.

Default – YES/NO – 2 classes

For more than 2 classes – Linear discriminant analisys

LDA – uses some approximations, assumes normal distribution AND SAME CORRELATION IN THE CLASSES

ROC curve – the closer it is to top left corner, the better the model, The line in the middle = 50% chance, below – worse, above – better.

True and false positive – take a look pg. 163

QDA – Quadratic discriminant analysis – ASSUMES DIFFERENT CORRELATION FOR DIFF. CLASSES

LOOK FOR CORRELATION BETWEEN X1/X2..Xn (predictors) within each class!!!

Roughly speaking, LDA tends to be a better bet than QDA if there are relatively few training observations and so reducing variance is crucial. In contrast, QDA is recommended if the training set is very large, so that the variance of the classifier is not a major concern, or if the assumption of a common covariance matrix for the K classes is clearly untenable.

KNN – FOR NON-LINEAR DECISION BOUNDRY, but no table of coefficients = hard interpretation

LOGISTIC REGRESSION AND LDA ARE SIMILAR, but LDA will perform better, if the distribution is Gaussian and common covariance matrix between classes. Conversely, if these are not met -> Log regression will perform better

QDA – quadratic – serves as a compromise between KNN and LDA and Log regr. –

Scale variables for KNN !!!!

LOG REGR in R = glm(family = binomial)