#1

AIC and BIC differ by the way they penalize the number of parameters of a model. More precisely, BIC criterion will induce a higher penalization for models with an intricate parametrization in comparison with AIC criterion.

#2

No, it isn’t. From the lecture 6 we known the F-test being using to determine accept or reject the H0, which tell us whether the predictor variables add any explanatory value to the model at all. The F-test with its associated p-value tells us that whether the linear regression model provides a better fit to the data than a model with no independent variables.. So in order to measure a model accuracy we should reference other statistic, like R^2.

#3

We can expand the data size, just like lecture video that we can use n=100000 instead of n=10.

#4

No, they’re not. Since we have proved the sum of residual is 0, This linear restriction on the residuals show that they cannot be independent. Then, Residuals depend on the fitted regression function, which depends on the same data that the residuals come from. With p fitting parameters, n residuals only have n – p degrees of freedom. The residual only independent when N being independent , normally-distributed random variable with mean 0 and variance a^2.

#5

From the textbook 5.3, if a mean function with one predictor X is smooth but not straight, integer powers of the predictors can be used to approximate E(Y|X).We know the data is correlated, but the relationship doesn’t look linear.

#6

One step ANVOA.

H0: All give rise to the same mean time to recovery from the viral illness

H1: At least one of the drugs give rise to a mean recovery time that is different from at least one of the others

H0: μ1=μ2=...μg.

H1: Not all μi are equal

We first load the data and create the relevant boxplot. The plot alone suggests a difference of means.

The aov() function is used to obtain the relevant sums of squares. Using the summary() function on the output from aov() creates the desired ANOVA table. And we can compare the p-value and significance level, if p < alpha,we will reject the null hypothesis and conclude at least one of the drugs give rise to a mean recovery time that is different from at least one of the others. If p < alpha