Conceptual Questions

Chapter 2 : Statistical Learning

Q1. For each of parts (a) through (d), indicate whether we would generally expect the performance of a flexible statistical learning method to be better or worse than an inflexible method. Justify your answer.

(a) The sample size n is extremely large, and the number of predictors p is small.
More flexible - worse

(b) The number of predictors p is extremely large, and the number of observations n is small.
(c) The relationship between the predictors and response is highly non-linear.

inflexible methods (like linear models) assume simple, restrictive functional forms for the relationship between the predictors and the response. If the true relationship given y f is highly non-linear, an inflexible method will fail to capture the true underlying pattern,which will lead to a high bias. Flexible methods avoid assumptions about the exact functional form of f, giving them the potential to accurately fit a wider range of possible shapes and substantially reduce bias, leading to a smaller overall test MSE making them a better overall fit

(d) The variance of the error terms, i.e. sig = Var(e), is extremely high.