1.(a) better

- 표본 크기가 클 때 flexible method 가 inflexible method보다 더 잘 맞는다.

1.(b) worse

- flexible method는 적은 관측치 값에 적합하지 않다.

1.(c) better

- 비선형적인 구조에서 더 자유롭고, 이 때 flexible method에 적합하다.

1.(d) worse

- 분산이 높으면 flexible method 에서 노이즈가 커진다.

2.(a) regression, inference

n = 500 firms in the US

p = profit, number of employees, industry

2.(b) classification, prediction

n = 20 similar products that were previously launched.

P = price charged for the product, marketing budget, competition price, and ten other variables.

2.(c) regression, prediction

n = weekly data for all of 2012

p = % change in the USD/Euro, the % change in the US market, the % change in the British market, and the % change in the German market.

6. Parametric은 모집단이 특정한 분포를 따른다는 가정하에 parameter를 추정하고 non-parametric 은 분포형태를 가정할 수 없는 경우에 분포형태에 대한 가정을 완화하여 parameter를 추정한다.

Regression 이나 classification 에서 parametric의 장점은 non-parametric에 비해 적은 obsevation이 필요하다.

그리고 단점은 가정된 f의 형식이 잘못되었을 때 f 를 부정확하게 추정하거나 flexible model을 사용할 때 과 적합할 수 있다.

7.(a) Obs.1 = 3

Obs.2 = 2

Obs.3 = √10

Obs.4 = √5

Obs.5 = √2

Obs.6 = √3

7.(b) green

Obs.5 가 K=1과 가장 가깝기 때문.

7.(c) red

OBs.2, 5, 6이 K=3과 가장 가깝고 여기서 2 = red, 5 = green, 6 = red이기 때문.

7.(d) small

K가 작을수록 non-linear decision에 유연하다. K가 커지면 좀더 linear decision에 유연해진다.