ISLR 4.7 Seungjun (Josh) Kim

4) a) If $x \in [0.05, 0.95)$, then observations we will use will be in $[x-0.05, x+0.05]$. There are two cases to consider :

i) $x < 0.05$ $[0, x+0.05]$ will be used for observations $\rightarrow (100x+5)\%$

ii) $x > 0.95$ $[x-0.05, 1]$ // $\rightarrow (105-100x)\%$

$\therefore$ Average fraction of observations used to make prediction

$$= \int_{0.05}^{0.95} 10\,dx + \int_0^{0.05} (100x+5)\,dx + \int_{0.95}^{1} (105-100x)\,dx$$

$$= [9.5 - 0.5] + [50x^2+5x]_0^{0.05} + [105x - 50x^2]_{0.95}^{1}$$

$$= 9 + (0.125+0.25) + (105 - 50 - 99.75 + 45.125) = 9 + 0.375 + 0.375 = 9.75\%$$

b) The fraction of observations used ~~to~~ make a prediction will be $0.0975 \times 0.0975 = 0.950625\%$

this holds only when X_1 & X_2 are independent from each other

c) For (b) above, for $p=2$, we did $(9.75\%)^2$.

Since we have $p=100$, $(9.75\%)^{100}$ = very small number close to 0%

d) If we have p features, the fraction of observations we use to make predictions becomes $(9.75\%)^p$ and this value becomes extremely small (close to 0) when p gets sufficiently large. Thus, this shows that there are not enough training observations near any given test observation.

e) This p dimensional hypercube contains, on average, 10% of training observations. Let the length of this hypercube denoted as l, then the number of observations this cube will contain can be expressed as l^p. Then, $10\% = 0.1 = l^p$

$$\therefore l = (0.1)^{1/p}$$

$\therefore$ (p=1) 0.1, (p=2) $0.1^{1/2}$, $l = 0.1^{1/100}$ (p=100)