VE572 Assignment 3

Liu Yihao 515370910207

2018-6-26

Qurestion 1

(a)

$$Pr(Y=1|X_1=x_1|X_2=x_2)=\frac{\exp(\beta_0+\beta_1x_1+\beta_2x_2)}{1+\exp(\beta_0+\beta_1x_1+\beta_2x_2)}$$

$$Pr(Y=1|X_1=40|X_2=3.5)=\frac{\exp(-6+0.05\cdot 40+3.5)}{1+\exp(-6+0.05\cdot 40+3.5)}\approx 0.378$$

(b)

$$\frac{\exp(-6+0.05x_1+3.5)}{1+\exp(-6+0.05x_1+3.5)}=0.5$$

$$\exp(-6+0.05x_1+3.5)=1$$

$$x_1=50$$

Question 2

(a)

$$odds(U.S.A|Y=1)=\frac{513}{487}$$

$$odds(China|Y=1)=\frac{116}{9}$$

(b)

$$OR=\frac{513/487}{928/72}=\frac{4617}{56492}\approx 8.173\times 10^{-2}$$

(c)

$$\beta_0=\ln\frac{116}{9}\approx 2.5819$$

$$\beta_0+\beta_1=\ln\frac{513}{487}$$

$$\beta_1=\ln\frac{4617}{56492}\approx -2.5043$$

(d)

$$\beta_1=\ln(OR)$$

(e)

$$\exp[\ln(OR) \pm 1.96 \times SE(\ln(OR))] = [0.0624, 0.1071]$$

Question 3

(a)

$$\int_0^{0.05} 0.05 + x dx + \int_{0.05}^{0.95} 0.1 dx + \int_{0.95}^{1} 0.1 - x dx = 9.75\%$$

(b)

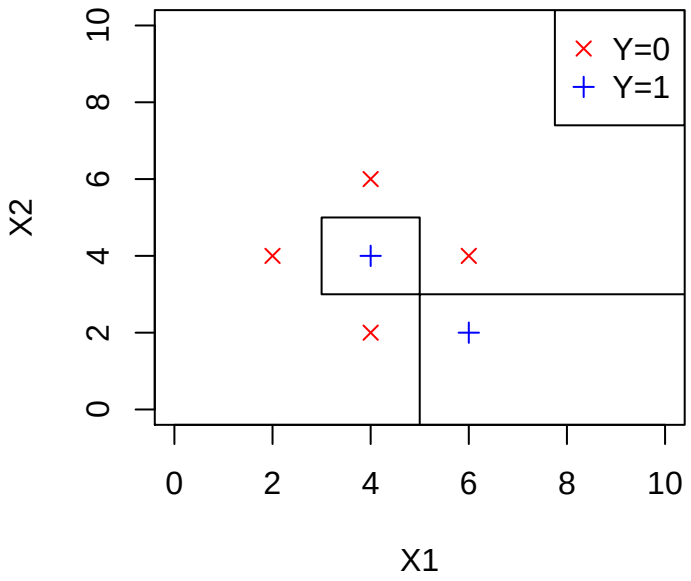
$$9.75\% \times 9.75\% = 0.950625\%$$

(c)

$$\sqrt[100]{0.1}$$

Question 4

(a)



(b)

$$Y(8, 1) = 1$$

Question 5

(a)

First, $C_1 = (-1.5, 0)$, $C_2 = (1.5, 0)$, so $(-2, 0.5)$ should be assigned to C_2 and $(2, -0.5)$ should be assigned to C_1.

Second, $C_1 = (-2.5, 0)$, $C_2 = (2.5, 0)$, so converges.

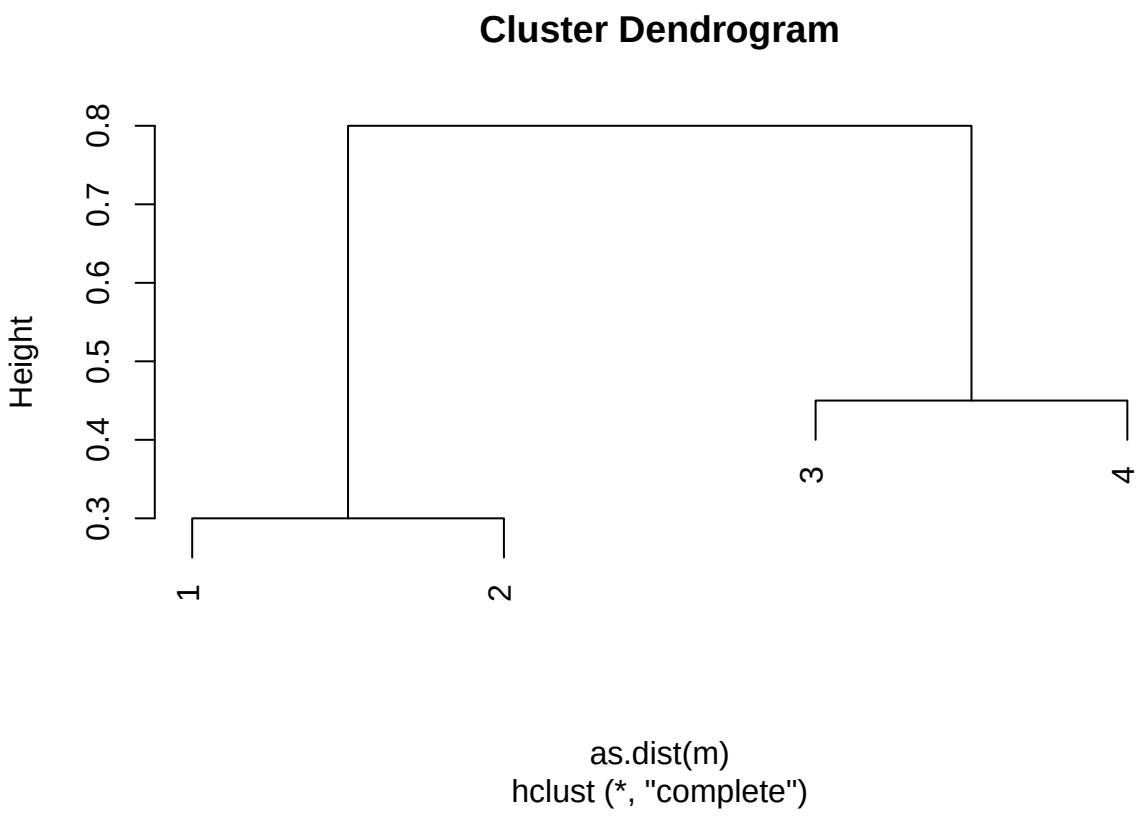
(b)

First, $C_1 = (0, 0)$, $C_2 = (0, 0)$, we can find that all points have the same distance to the two centersm so converges.

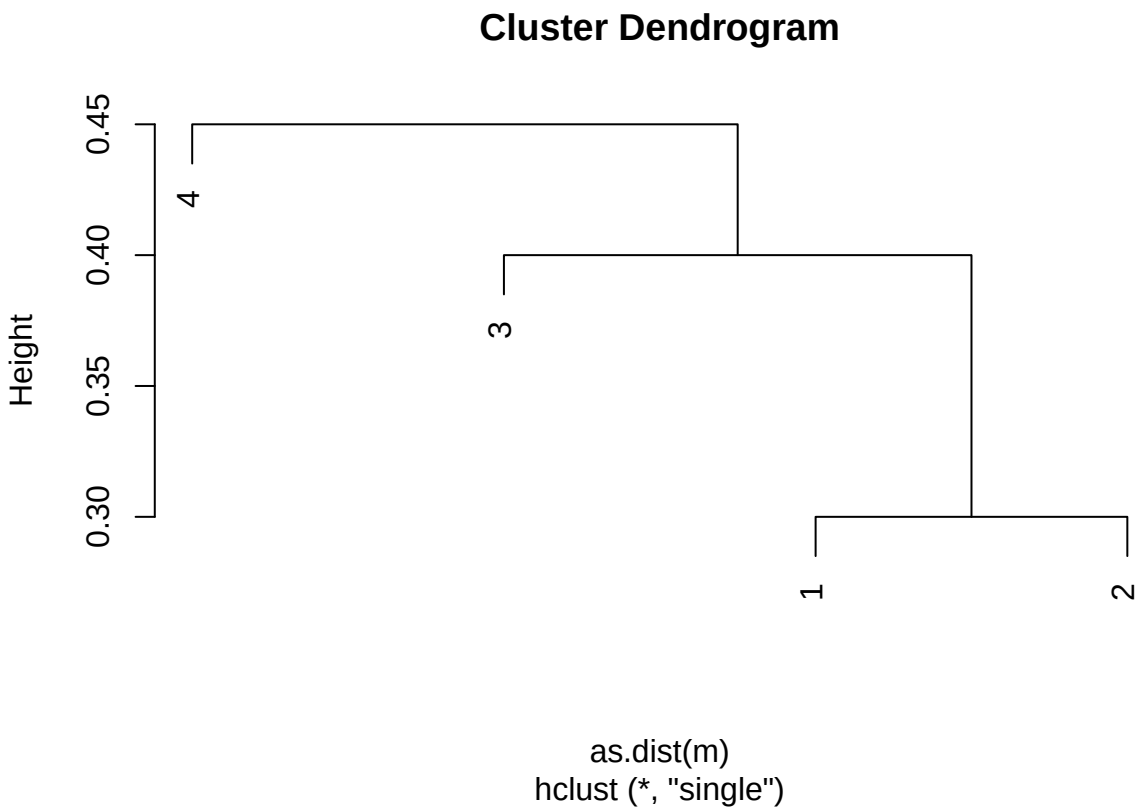
I notice that different initialization can give very different results in K-means clustering, so the randomness of initialization is very important.

Question 6

(a)



(b)



(c)

Observations 1, 2 in C_1 and 3, 4 in C_2

(d)

Observations 1, 2, 3 in C_1 and 4 in C_2