Statistical learning assignment 8- chapter 4

孫浩哲　　M072040002

November 8, 2018

7.

$$\begin{aligned} p_{yes}(x) &= \frac{\pi_{yes}\exp(-\frac{1}{2\sigma^2}(x-10)^2)}{\pi_{yes}\exp(-\frac{1}{2\sigma^2}(x-10)^2)+\pi_{no}\exp(-\frac{1}{2\sigma^2}x^2)} \\ &= \frac{0.8\exp(-\frac{1}{72}(x-10)^2)}{0.8\exp(-\frac{1}{72}(x-10)^2)+0.2\exp(-\frac{1}{72}x^2)} \end{aligned}$$

$$\Rightarrow p_{yes}(4) \approx 0.752$$

Approximately 75.2% that the company will issue a dividend.

8.

The training error rate is close to 0 if $K = 1$, and average error rate is 18%, that means error rate of 36% on the test data.So we prefer logistic regression.

9.

(a.)

$$\frac{P(X)}{1-P(X)} = 0.37 \Rightarrow 1.37P(X) = 0.37, P(X) \approx 0.27$$

(b.)

$$P(X) = 0.16$$

$$odds = \frac{P(X)}{1-P(X)} = \frac{0.16}{0.84} \approx 0.19$$

12. (a)

```
> power=function()
+ {
+   result=2^3
+   print(result)
+ }
> power()
[1] 8
```

(b)

```
> power2=function(x,a)
+ {
+   result=x^a
+   print(result)
+ }
> power2(3,8)
[1] 6561
```

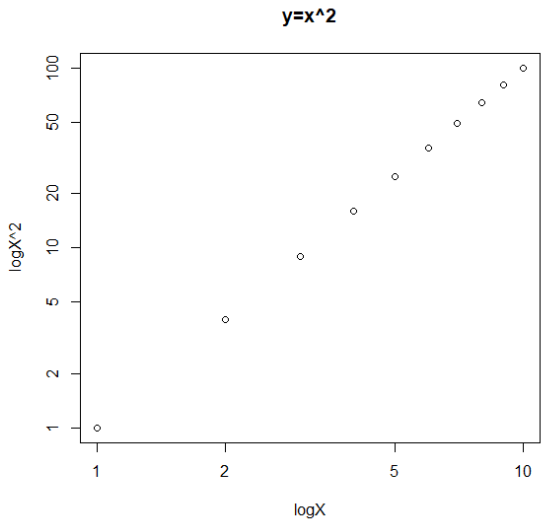
(c)

```
> power2(8,17)
[1] 2.2518e+15
```

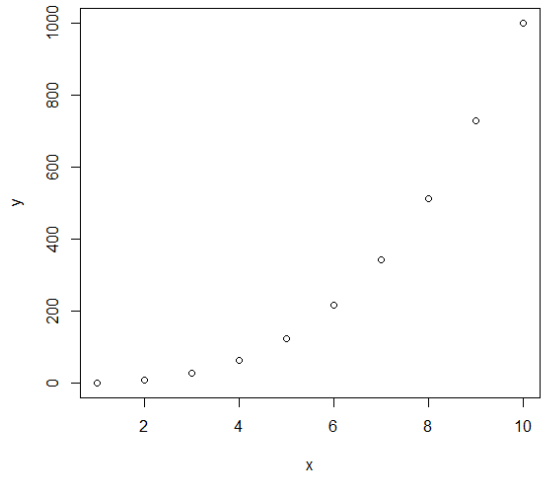
(d)

```
> power3=function(x,a)
+ {
+   result=x^a
+   return(result)
+ }
```

(e)



(f)



13.
```
> #######logistic regression#######
> {
+ class=rep(0,nrow(Boston))
+ class[which(crim>median(crim))]=1
+ boston=data.frame(Boston,class)
```

```
+ attach(boston)
+ model=glm(class~.-crim,data=boston,family=binomial)
+ tr=boston[1:400,]#training set
+ te=boston[401:506,]#test set
+ tecr=te$class#real test
+ pro=predict(model,data.frame(te),type="response")
+ tepre=rep(0,nrow(te))
+ tepre[which(pro>0.5)]=1
+ mean(tepre!=tecr)#test error
+ }
[1] 0.04716981
> #######LDA#######
> {
+ LD=lda(class~.-crim,data=boston,family=binomial)
+ telda=predict(LD,data.frame(te))
+ mean(telda$class!=tecr)
+ }
[1] 0.0754717
> #######KNN#######
> {
+ kn=knn(tr,te,tr$class,k=1)
+ mean(kn!=tecr)
+ }
[1] 0.0754717
> {
+ kn=knn(tr,te,tr$class,k=10)
+ mean(kn!=tecr)
+ }
[1] 0.0754717
> {
+ kn=knn(tr,te,tr$class,k=100)
+ mean(kn!=tecr)
+ }
[1] 0.09433962
```

In this case we can find that the logistic regression is the best approach to predict, and the test error of LDA is equal to KNN when $K = 1, 10$.